University of Dundee

MASTER OF SCIENCE

A Pilot Study- Identify Genetic Variants for Diabetic Cataract Using GoDARTS Dataset

Zhang, Kaida

Award date:
2016

A Pilot Study- Identify Genetic Variants for Diabetic Cataract Using GoDARTS Dataset

Kaida Zhang

A dissertation submitted for the degree of Master of Science by Research from the University of Dundee

School of Medicine

University of Dundee

September 2016

Abstract

Background: Diabetic cataract is one of the major eye complications of diabetes. It was reported that cataract occurs two to five times more frequently in patients with diabetes compared with those with no diabetes. The purpose of this study was to identify genetic contributors of diabetic cataract based on a genome-wide association approach using a well-defined Scottish diabetic cohort.

Methods: A diabetic cataract case in this study was defined as a type 2 diabetic patient who has ever been recorded in the linked e-health records to have cataracts in one or both eyes and who had previous cataract extraction surgeries in at least one eye. A control in this study was defined as a type 2 diabetic individual who has never been diagnosed as cataract in the linked e-health records and had no history of cataract surgeries. A standard genome-wide association approach was applied. Besides, the logistic regression was used to analyze the potential risk factors including Age, Gender, Body Mass Index(BMI), Alcohol intake, Total serum cholesterol, High-density lipoprotein(HDL)-cholesterol, Low-density lipoprotein(LDL)-cholesterol, Blood pressure, HbA1c and Serum triglycerides chosen from the literature review.

Results: Overall, we have 1986 diabetic cataract cases and 3429 controls in the genetics of diabetes audit and research in Tayside Scotland (GoDARTS) dataset. We set the significant *P* value of Single Nucleotide Polymorphisms (SNP) in the project as 10^{-6}, there are 7 associated Single Nucleotide Polymorphisms in the range of genome-wide significance we set, including rs10197646 (*P* $4.12x10^{-07}$), chr13:48026216:D (*P* $4.15x10^{-07}$), rs7582173 (*P* $4.30x10^{-07}$), rs62168795 (*P* $5.59x10^{-07}$), rs1381015 (*P* $7.12x10^{-07}$), rs2269547 (*P* $7.25x10^{-07}$), rs523355 (*P* $8.63x10^{-07}$). The age-adjusted prevalence of diabetic cataract was 24.9% in the Tayside. We also identified age (odd ratio [OR] 0.955, 95%confidence interval [CI] 0.948-0.962), female (OR 1.191, 95%CI 1.055-1.345), systolic blood pressure (OR 0.997, 95% CI 0.994-0.999) diastolic blood pressure (OR 1.004, 95% 1.001-1.008), current smoker (OR 1.313, 95% CI 1.034-1.667), BMI in 2nd Quartile 27.71-31.32 (OR 0.838, 95% CI 0.703-0.998), total serum cholesterol in 2nd Quartile 3.92-4.37 (OR 0.798, 95% CI 0.642-0.992), serum HDL cholesterol in 3rd Quartile 1.36-1.48 (OR 0.737, 95% CI 0.596-0.910), and serum triglycerides in 3rd Quartile 2.24-2.40 (OR 0.393, 95% CI 0.316-0.490) as associated significant factors with diabetic cataract in Scottish population.

Conclusions: We identified the 7 significant SNPs related with the potential genes in Tayside population and found supporting evidence that *MAP3K19, R3HDM1*, *GGA1*, *CCT7* genes are associated with diabetic cataract. The role of genes in the cataractogenesis needs to be reevaluated in future studies. The risk and protective factors were identified with GoDARTS dataset.

Contents

List of Figures

List of Tables

Acknowledge

This project was supported by National Health Service Tayside, School of Medicine and Population Health Sciences

I would like to give my sincere gratitude to my supervisor Dr. Weihua Meng who gave me many important helps, taught me many knowledge related to my project and provided me kind encouragement. And I would like to thank Dr. Helen C Looker and Dr. Fiona LR Williams who also gave me a great support and many suggestions. Then I would like to express my gratitude to Miss. Emma Carroll who helped me during my project. I am pleased to acknowledge Prof. Colin Palmer who supported this project. In addition, I would like to give many thanks to my parents and my girlfriend Miss Chong Chen who supported me all the time.

At last, I would like to express my gratitude to all those who helped me during my project.

Declaration

I hereby declare that I am the author of this dissertation, it is a record of the work that has been done by me, and it has been accepted for a higher degree. I also state that all references cited have been consulted by me and the conditions of the relevant ordinance and regulations have been fulfilled.

Signed……..

Date……..

(Student Name)

Signed……..

Date……..

(Supervisor)

Chapter 1. Introduction

1.1. Concept of Cataract in Public Health and Epidemiology

Public health is usually described as "the science and art of preventing disease, prolonging life and promoting health through organized efforts and informed choices of society, organizations, public and private, communities and individuals (1). The focus of public health is to improve the health and quality of life in human beings through the prevention and treatment of disease and other physical and mental health conditions. According to the latest reports from the World Health Organization (WHO), cataract is becoming a major public health concern across the world (2), is increasing in prevalence, is affecting quality of life, and is considered a burden on the health care system. Especially for diabetic patients, cataract is a major cause of blindness in developed and developing countries (3).

In the public health field, epidemiology is the cornerstone of identifying risk factors for disease and targets for preventive healthcare, which develop methodology used in clinical research, public health studies, and even basic research in the biological sciences (4). With the increase in challenging health problems all over the world, genetic epidemiology, known as “a science which deals with the etiology, distribution, and control of disease in groups of relatives and with inherited causes of disease in populations”, has developed as one of the most important studies in recent years for determining the disease and health in families and populations (5). With the help of advances in genotyping technology, it is now feasible to conduct large-scale Genome-Wide Association Studies (GWAS) that genotype many thousands of single nucleotide polymorphisms (SNP) in thousands of individuals. These have led to the discovery of many genetic polymorphisms that influence the risk of developing many common diseases. As we know, there are many studies about cataract which focus on its distribution, determinants and risk-exposure associations. Moreover, genetic variants are considered as one of the determinants in cataract research, which now extend to a new field to explore the mechanism of diabetic

cataract.

1.2. Concept of Cataract and Classification

Cataract is clouding of the lens of the eye which prevents clear vision (6), it is also described as the opacity of crystalline lens in the eye (7). In the process of cataract studying, at least four systems have been developed to classify and grade lens opacities. These include the grading system developed at the Wilmer Ophthalmological Institute of Johns Hopkins University, the Oxford Clinical Cataract Classification and Grading System(8), the Lens Opacity Classification System (LOCS) (9), and the Wisconsin system(10). These involved photographic documentation of the lens, using standard camera and flash settings, with interpretation of the lens photographs undertaken by trained graders.

1.3. Diagnosis

To diagnose the cataract, detailed visual history and a full medical history should be taken, and the impact of cataract on the patient's lifestyle should be evaluated. Questionnaires can be helpful in describing symptoms but should be used in conjunction with the patient's history and examination when deciding on surgery. Following this, a complete ophthalmic examination should be performed by ophthalmologists (11) (Figure 1.1).

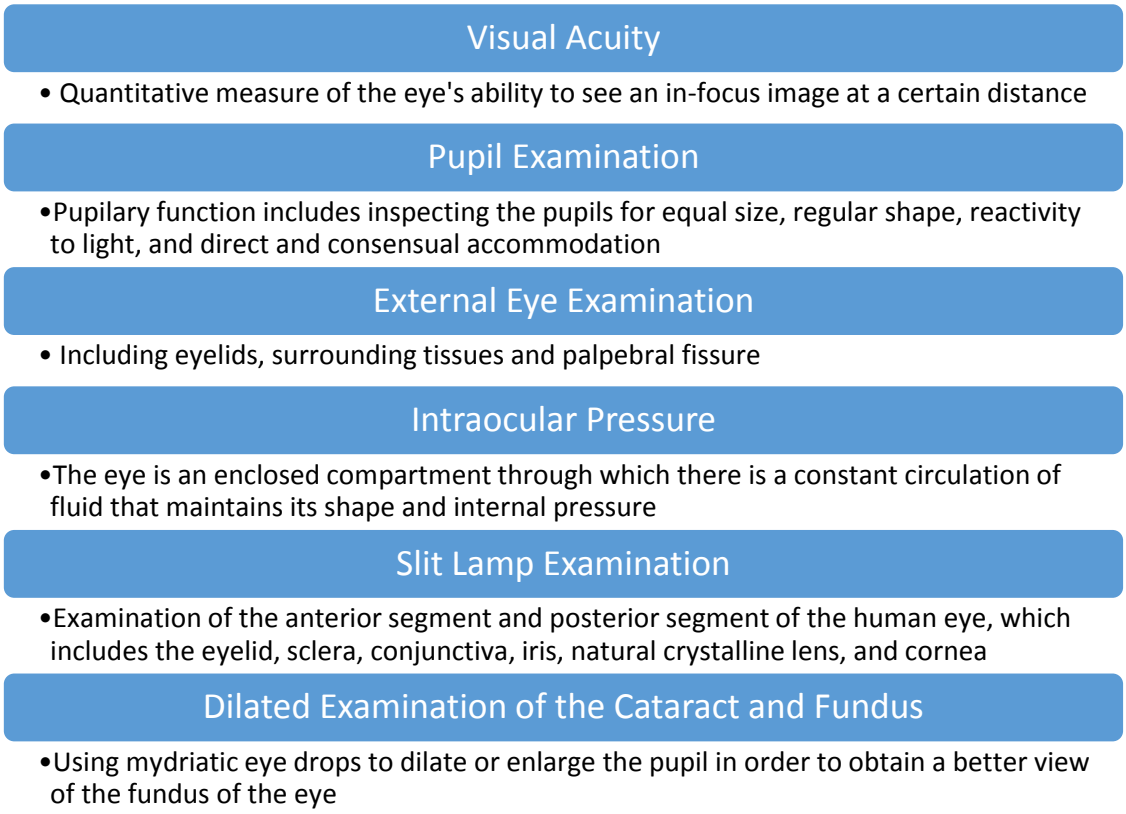

Figure 1.1: Stages of a full cataract examination

1.4. Classification

Cataract can be classified as nuclear, cortical or posterior by using the lens opacities classification system LOCS III (9).

- Nuclear cataract is the most common type of cataract and involves the central or nuclear part of the lens; this type of cataract often results in a shift to nearsightedness and causes visual problems.

- Cortical cataracts are due to the lens in the outer layer changing to opaque, which occurs in the periphery of the lens causing fissuring. These kinds of cataracts are detected by using an ophthalmoscope, or other magnification system, and the appearance is often described as the white spokes of a wheel.

Posterior capsular cataracts, defined as cloudy at the back of the lens adjacent to the capsule in which the lens sits, are due to strong light focusing on the back of the lens, which can cause disproportionate symptoms for their size. This type of cataract is proven to be the most common type in diabetic patients (12).

According to previous clinical and epidemiological studies between diabetes and cataracts, cataract is considered a major cause of visual impairment in diabetic patients and the progression of cataract is elevated in patients with diabetes mellitus (13, 14). Diabetes mellitus is a systemic condition affecting numerous organs, including the eye, which has serious influence on the development and progression of ocular complications in diabetic patients. The International Diabetes Federation offered evidence that more than 285 million people suffer from diabetes mellitus worldwide; this number is expected to increase to 439 million by 2030 (3). Both cataract and diabetes are a big challenge in the public health field and the economy in developing and developed countries. As this is the case, it is recommended that an emphasis is put on looking for the factors that influence cataract in diabetic patients.

1.5. The Mechanism of Cataract

The eye, as part of our visual system, contains a lens. The lens is a clear structure behind the pupil and iris. The lens focuses light on the retina which is the back part of the eye that sends sight signals to the brain. When the lens becomes cloudy, vision blurs; we call this a cataract (15) (Figure 1.2).

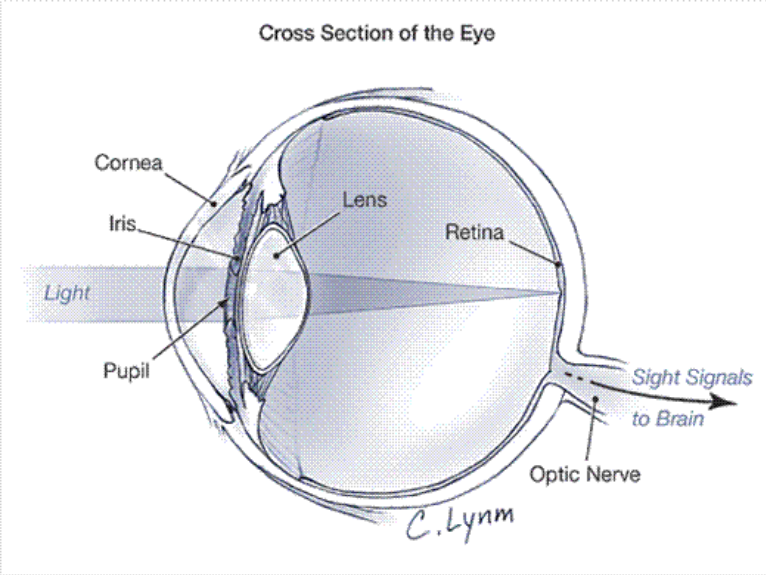

Figure 1.2: Cross section of the eye

The lens is an avascular organ where nutrients and oxygen are supplied by a blood-aqueous barrier, formed by the non-pigmented layer of the epithelium of the ciliary body and the endothelium of the blood vessels of the iris, (16) and which depends strongly on glucose metabolism. Therefore, for diabetic patients, hyperglycaemia-initiated pathogenethy mechanisms perform an essential role in diabetes-associated changes in lens metabolism and cataract formation, which include, but are not limited to, increased aldose reductase (AR) activity, non-enzymatic glycation/glycooxidation and oxidative-nitrosative stress (16). In fact, there have been a number of new pharmacological approaches supporting the breakthroughs in understanding the mechanisms of diabetic cataractogenesis.

The process of lens glucose uptake is insulin-independent, which has different reactions in diabetes. Under non-diabetic conditions, glucose is phosphorylated by hexokinase and metabolized by glycolysis and pentose phosphate pathway. In diabetes, hexokinase has become saturated (17) and excesses of glucose enters the sorbitol pathway (17, 18) with two reactions including aldose reductase (AR) catalyses NADPH-dependent reduction of glucose to sorbitol and sorbitol dehydrogenase SDH catalyses NAD-dependent oxidation of sorbitol to fructose. Especially in diabetic cataract formation (19, 20), the increased

AR through the sorbitol pathway activity in the lens causes intracellular sorbitol accumulation, osmotic stress, and a decrease in free cytosolic NAD+/NADH+ ratio (17, 21, 22); all consequently contribute to multiple metabolic and signal transduction changes, and affects transcriptional regulation and gene expression in the lens fibers and tissue sites for diabetic cataract (23, 24). Results of a previous animal study reported that the wild type Streptozotocin (STZ), when examined in diabetic mice, has a very low AR expression in the lens, and does not lead to sorbitol accumulation and lens opacification (18, 20); on the contrary, cataract developed in diabetic and galactose-fed mice expressing human AR in the lens (20). Furthermore, two reports demonstrated that sorbitol accumulation and the rate of cataractogenesis were greater in diabetic AR expressing mice (25, 26). All of this evidence indicates that increased AR activity plays a key role in diabetic cataract formation. Recent studies have shown the increased AR activity is likely to influence the development of diabetic cataract through osmotic stress. An extensive research in the crystalline lens by Kinoshita (21) suggests that osmotic stress is important in the development of fast cataractogesis in young diabetic and galactose-fed animal models. However, the increased AR activity might have greater important influence on slow cataract development in mature STZ-diabetic rats treated with insulin (27, 28) and diabetic patients. Although the AR activity in the human lens is rather low (29), it is unknown exactly how much of the enzyme is necessary to produce a detrimental metabolic consequence that includes cataract. Still, the important role of AR in the progress of diabetic cataract in humans has been proven by a genetic study of cataract risk factors in Hong Kong Chinese patients (30).

Non-enzymatic glycation and glycoxidation have important implications in the pathogenesis of diabetic complications (16). The test found advanced glycation end products (AGE) in the process of non-enzymatic glycation and glycoxidation accumulate in the lenses of some types of diabetic animals, as well as cataractous lenses in diabetic patients (31). The AGE formation often occurs in both lens epithelial (32) and fiber cells (33-35), and the AGE-modified β- and γ-crystallins have been detected in strepotozotocin

diabetic rats (34, 35). Similarly, the AGE accumulation has also been identified in cataractous lenses in humans with diabetes (36, 37). The AGE accumulation contributes to the human lens protein aggregation and subsequent insolubilzation (38), where the high and low molecular weight crystallins andα-crystallin from diabetic displayed elevated furosine content (36), and the AGE-linked autofluorescence has been reported to an increase in cataractous lenses of diabetic subjects when compared with the non-diabetic controls (37). Moreover, the modification of α-crystallin caused multiple changes in lens including structural changes, cross-linking, coloration and subsequent insolubilization leading to a scatter of light (39). Moreover, a follow-up study demonstrated the modification of α-crystallin might cause unfolding and decreased stability leading to enhanced proteolysis in the human lens (39). All of these studies tend to support the influence of AGE accumulation during Non-enzymatic glycation and glycoxidation on the development of diabetic cataract. At the same time, there are other experimental and clinical studies reporting that an inhibitor of AGE in diabetic rats was ineffective in preventing fast cataractogenesis (40). Although a number of researches and studies have shown the important influence of Non-enzymatic glycation and glycoxidation with AGE accumulation on the pathogenesis of diabetic cataract, it is still a subject of controversy.

What is more, in several diabetic complications, the oxidative–nitrosative stress was recognized as a key mechanism in the pathogenesis (17). Oxidative stress presents in the early stages of diabetes in the lens, and is manifested by the accumulation of lipid peroxidation products, reduced glutathione (GSH) and other metabolites. Oxidative-nitrosative stress also causes an interaction with multiple mechanisms including increased AR activity, non-enzymatic glycation and glycoxidation, and elevation of cytosolic Ca 2+; which together have influences on the development of diabetic cataract (41). Using transgenic mice that over express aldose reductase (AR) in their lenses, studies have found that the flux of glucose through the polyol pathway is the major cause of hyperglycemic oxidative stress in this tissue (42). AR reduction of glucose to sorbitol probably contributes to oxidative stress by depleting its cofactor NADPH, which is also

required for the regeneration of GSH. Sorbitol dehydrogenase, the second enzyme in the polyol pathway that converts sorbitol to fructose (19, 20), also contributes to oxidative stress, most likely because depletion of its cofactor NAD+ leads to more glucose being channeled through the polyol pathway. There are few studies presenting clear evidence of complete correction of oxidative stress-nitrosative mechanism in diabetic cataract patients (43, 44), where chronic oxidative stress generated by the polyol pathway is likely to be an important contributing factor in the slow-developing diabetic cataract, as well as in the development of other diabetic complications (42).

1.6. Genetic Factors about Diabetic Cataract

Based on genetic epidemiology, Genome-Wide Association Studies (GWAS) can be used to compare the DNA of participants having varying phenotypes for a particular disease. Participants in a GWAS study are treated as cases and controls. This approach is known as phenotype-first, in which the participants are classified first by their clinical manifestations, as opposed to genotype-first. Each person gives a sample of DNA, from which millions of genetic variants are read using SNP arrays. If one type of the variant (one allele) is more frequent in people with the disease, the variant is said to be associated with the disease. The associated SNPs are then considered to mark a region of the human genome that may influence the risk of disease (45, 46).

1.6.1. Genetic Variants in Cataract

Understanding the genetic factors related to the diabetic cataract would assist in identifying the potential causal mechanisms for future researches. Animal models have been widely applied in genetic research in cataracts. A small scale Genome-Wide Associated Study (GWAS) proposed 15 loci genome related to diabetic cataract in the Chinese population; the loci of which contain candidate genes of PPARD, CCDC102A, GBA3, NEDD9, GABRR1/2, RPS6KA2, tcag7.1163, TAC1, GALNTL1 and KIAA1671 (47). These candidates involved the relative mechanisms of cataract formation, and GWAS is considered as a useful and efficient method to identify potential candidate genes

for common complex disorders using DNA chips (48). The DNA chips can genotype thousands of single nucleotide polymorphisms (SNP) in individuals to compare variants between cases and controls.

1.6.2. Genetic Variants in Diabetes

Genome-Wide Association Studies have ascertained these genetic variants to be associated with an increased risk of diabetes. These studies concluded the Transcription Factor 7-like 2 (TCF7L2) has the strongest susceptibility effect (49), which had an influence on the development of Type 2 diabetes via its role in the development of pancreatic islet cells including beta-cell survival and secretory function; development of myocytes and adipocytes. Also, other genes discovered from the GWAS include HHEX controlling beta islet cell development; WFS1 is responsible for the survival and function of the beta islet cells; SLC30A8 affects insulin production and secretion through the transport of zinc to the pancreatic beta islet cells; GIPR is associated with reducing beta cell function; KCNJ11 blocks the release of serum glucose to cells, and IGF2BP2 is linked to increased adipose levels and insulin resistance.

1.7. Epidemiological Risk Factors for Cataract

The development of cataract is a complex process affected by multiple risk factors. To understand this better, the association between risk factors and cataract have to be explored. Therefore, the following table concluded the potential risk factors which could affect the development of diabetic cataract according to the previous studies and researches (Table 1.1).

1.7.1. Personal Characteristics

a) Age

Epidemiological studies on lens opacities mostly focus on age-related cataract; the development of cataract is a continuous process and some opacification occurs systematically with age (50). In the UK, there are a number of population based studies

for prevalence of cataract. According to the North London Eye Study data, the prevalence of visually impairing cataract rose steadily with age: 16% in the 65 to 69 year age group, 24% in people 70 to 74 years of age, 42% in those 75 to 79 years of age, 59% in those 80 to 84 years, and 71% in people aged 85 years old or over (51). Another study, The Lens Opacities Case-Control Study, explored the risk factors for age-related nuclear, cortical, posterior subcapsular and mixed cataracts, and showed cortical cataracts were associated with age with relative risk (RR) = 11.4 for age 70 years compared to age 50 years, nuclear cataracts have 38.6 times more risk of developing in those aged 70 years when compared to those aged 50 years (52). Even taking gender into consideration, the prevalence of cataract is still significantly associated with increasing age, the increase is 3-fold for males and approximately double for females between the ages of 60-64 years and 75 years and older, as shown in a Nepalese blindness survey (53). Similarly, an American Health and Nutrition Examination Survey showed the prevalence of cataract in a group aged 45-64 years is three times higher than that for the 65-75 year age group, for both females and males (54). Therefore, age is a major risk factor in the development of cataract.

b) Gender

If taking adjustment for other risk factors, women would commonly show a higher cataract prevalence when compared to men (50). Many studies worldwide have reported a higher prevalence of cataract among women (55, 56), although in some studies this varied by the type of opacity (57, 58). The Blue Mountains Eye study released in Australia showed 53.3% of cataract prevalence for women and 49.7% for men in moderate or advanced nuclear opacities. The cortical cataract was present in 25.9% of women and 21.1% of men, and posterior subcapsular cataract was less frequent (59); which is consistent with three other studies. The estimate of age-adjusted prevalence in the United States for women was significantly higher than men (55), and in a case-control study to find the risk factor for lens opacity females had a higher prevalence than males in the different cataract classification groups (60). What is more, a longitudinal study also confirmed the higher relative risk of lens opacity in women when compared with men

(61).

c) Race

The influence of the human race on development of cataract is still controversial in many studies. The Singapore Epidemiology of Eye Diseases Study concluded different Asian ethnicities had a higher prevalence and earlier age of onset of cataract than Europeans (62). The Baltimore Eye Survey stated cataract accounted for 27% of blindness among blacks, among whom it was four times more common than among whites, and whites were almost 50% more likely than blacks to have cataract extraction (63). The results from a Longitudinal Study of Cataract show the 2.94 attributable risk in increased nuclear opacification for whites (61). On the other hand, the National Health and Nutrition Examination Survey found the cortical cataracts had 3.5 relative risk (RR) in those of black race when compared with those of white race (52).

d) Social Economics and Education

A number of studies have identified the social economic status and level of education as important risk factors for cataract, where multivariate analyses demonstrated that cataracts were more severe when the median income was lower and unemployment rate was higher (64). However, they are probably confounded by factors that were omitted or that are not yet understood. Therefore, a case-control study of cataract risk factors has now been performed to support that low education and no education were associated with an increased risk of cataract after adjustment for age and other demographics factors (65). Another case-control confirmed that education and income were associated with age-standardized cataract prevalence for certain Asian subgroups (62).

1.7.2. Medical History

a) Diabetes

Diabetes was strongly associated with the risk of all types of cataract. Biochemical studies of cataractous lenses in those with diabetes and galactosemia showed abnormalities in the

levels of electrolytes, glutathione, glucose or galactose (66); which can lead to hyperosmotic effects in the lens such as lens fiber swelling, vacuole formation and opacification (14, 66). A number of epidemiological studies have supported this significant association between cataract and diabetes. The Framingham Eye Study observed the association of increasing blood sugar level with cataract (14), and several clinic-based studies have also found diabetes to be a risk factor for cataract (67-69). A population-based study, conducted in South Central Wisconsin, identified the important characteristics for diabetic cataract patients; for younger-on-set diabetics, the most important factor was increasing duration of diabetes, for older-onset diabetics it was age at the time of the survey. The National Health and Nutrition Examination Survey analyzed a three- to four-fold excess risk of senile cataracts for diabetic individuals younger than 65 years old, but the risk did not persist for those over 65 years of age with the Framingham data (70).

b) Smoking

For smoking, evidence shows a causal relationship between cigarette smoking and cataract (71), which has also been demonstrated by study testing the dose-effect relationship between a pack of cigarettes smoked and the opacification degree, and also shows that increased duration of smoking is related to the risk of cataract (72, 73). The literature reviews suggest a strong association between the development of nuclear cataract and smoking, but there is limited evidence for the risk of cortical and posterior subcapsular cataract (74). Smoking may cause damage to the lens by increasing oxidative stress, by lowering levels of circulating antioxidants or by increasing lens cadmium levels (75-77).

c) Ultraviolet Radiation

Lens opacification has been related to ocular exposure to ultraviolet radiation, particularly ultraviolet B (UV-B), because the lens has the ability to absorb UV-B and UV-A (78). The animal experiments claimed that the changes in lens clarity were related to different

intensity exposures and chronic exposure to UV-B (79); these findings have also been addressed in several epidemiological studies. For example, the National Trachoma and Eye Health Programme in Australia revealed a dose-response relationship of increasing prevalence of cataract with increasing levels of ultraviolet B radiation (80). In Nepal, another national survey found a positive correlation between cataract prevalence and average daily sunlight hours (81). In some areas of Nepal, there is a daily average of 12 hours of sun exposure which results in a prevalence of cataract nearly four times higher than those living with seven hours of exposure. In the United States, the National Health and Nutrition Examination survey analyzed data from almost 10,000 eye examinations in 35 geographic locations to conclude that people living with higher total annual sunshine hours had a higher prevalence of cataract (82).

d) Blood Pressure/Hypertension

In the Framingham Eye Study, systolic blood pressure was significantly higher in those with cataract when compared to non-cataract individuals in the same age and sex groups (14). A National Health and Nutrition Examination Survey showed a twice increased risk for having Posterior subcapsular cataract in those with systolic blood pressure of 160 mmHg compared to those with 120 mmHg (52). The documented case-control study between India and the U.S. suggested an increased risk for nuclear and mixed cataracts for each 20 mmHg increase in systolic blood pressure (83). However, the reason for blood pressure potentially affecting cataract development is still unclear, and more studies are needed to find the accurate mechanism by examining the independent effects of blood pressure and use of antihypertensive medications (84, 85). There was an investigation which showed that cataract could be prevented by both acute and chronic dietary sodium restriction on genetically hypertensive, salt-sensitive rats; suggesting a stronger role for extracellular fluid volume status than that of sustained arterial hypertension in cataract development (86).

1.7.3. Biochemical Variables

Previous experimental studies determined the lipid composition in protein fractions of human lens and senile cataracts (87), suggesting cholesterol is the major lipid component of the ocular lens (88), and the increased risk of senile cataracts related to cholesterol/phospholipid ratio (87). Some epidemiological studies report the different levels of total serum cholesterol, LDL cholesterol and HDL cholesterol are related to the cataract (89, 90). Also, experimental research has found some inhibitors of cholesterol can block cholesterol accumulation by these lenses and can produce cataracts in dogs (91), moreover, patients using cholesterol-lowering statin drugs may be at an increased risk of developing age-related cataracts (92). Based on the updated literature review, plasma retinol level was recognized as a significant risk factor for nuclear and mixed cataracts, as well as cataract surgery, which strongly suggest that vitamin A may be a protective for cataract (93, 94). Two other studies on plasma retinol (95, 96) gave inconsistent results; the cataract retinol relation could be confounded possibly by plasma carotenoid, which may decrease the risk of cataracts severe enough to require extraction (97).

Indeed, apart from serum cholesterol and Vitamin A, Serum triglycerides, HbA1c and urinary protein, all have some kind of influence on the progress of cataract (98).

Previously, one of researches to quantitatively evaluate the prevalence and risk factors of cataracts in Korean patients with type 2 diabetes mellitus discussed the risk factors like sex, age, duration of diabetes, fasting blood sugar, HbA1c, creatinine, and total cholesterol between patients with and without cataracts. Moreover, another cross-sectional study used the LOCS III to classify the cataract as different subtypes, and revealed that the age- and sex-adjusted prevalence of cataract in the study was 65.7% (95% CI, 65.6-65.8). Mixed cataracts were more common than monotype ones (41.6% vs. 19.4%). The prevalence of cataract was higher in women, subjects with known diabetes and those with longer duration of diabetes (51.4%, 50.3%, and 64.5%, respectively). The risk factors for any type of cataract were increasing age (OR, 1.14; 95% CI, 1.11-1.16).

Therefore, we concluded the table 1.1 to demonstrate the potential risk factors for the diabetic cataract from recent studies.

Table 1.1: Risk factors for cataract

	Risk factor	Higher risk/comment
Personal Character	Age	Increasing age
	Gender	Female
	Race	Controversial
	Social economics	Significant
	Education	Significant
Medical History	Diabetes	Significant
	Smoking	Risk factor
	Ultraviolet Radiation	Risk factor
	Blood Pressure/Hypertension	significant
Biochemical variables	Cholesterol	Significant
	Plasma retinol	Controversial
	Triglycerides	
	HbA1c	
	Urinary protein	

1.8. Epidemiology of Cataract

Cataract epidemiology is becoming an intensive research. Over the last decades, several large population based studies have offered new results on the prevalence of cataract in many regions, which has become more meaningful as more attention is directed to improving the assessment and measurement of both cataract and the potential risk factors. The reports for the estimate of blindness and visual impairment in 2010 show there are 285 million affected by visual impairment or blindness. According to data from the WHO, cataract is the leading cause of visual impairment (33%) and blindness (41%) (99).

1.8.1. The Epidemiology of Cataract in Asia

As the world's largest continent, with more than half of the world's population, up to 20 million Asians are estimated to be blind by the WHO (100). However, the prevalence of cataract between studies varies widely, and the comparison is confounded by differing population characteristics and diagnostic methods of lens opacity. Mostly based on the clinical examination Lens Opacity Classification III (LOCS III), the Tanjion Pagar Survey showed the prevalence of cataract for people over 40 years in Singapore is 34.7% (101), compared with a cataract prevalence of 61.9% in Indian people from the same age group in the Araind Eye Study (57). Shihpai Eye Study for Taiwan people over 65 years demonstrated the prevalence of 59.2% (102), and the Sumatra Eye Study in Indonesia offered 23.0% people suffering from cataract (103). However, there are few data on cataract surgery rates in Asia. The WHO suggested that an annual rate of 350 surgeries per 100,000 is a useful target to tackle the burden of cataract blindness (104); this threshold may have been achieved in some Asian countries (105).

1.8.2. The Epidemiology of Cataract in America

Institutes in the United States have conducted many population studies and have more accurate data for cataract. According to cataract reports from the National Eye Institute in 2010 (106), white Americans aged 40 and older had the highest prevalence rate of cataract (18%) followed by black Americans (13%); Hispanic Americans had the lowest rate of cataract (12%). Among all people with cataract in the U.S., the vast majority were white (80%) with lower rates for black people (8%) and Hispanic people (7%). Cataract also varied by gender, 61% of Americans with cataract were women with 39% being men. This report also indicated that the number of people in the U.S. with cataract is expected to double from 24.4 million to approximately 50 million. The majority of cases will affect white people; however Hispanic Americans are expected to have the most rapid increase in prevalence from 1.76 million cases to 9.51 million.

1.8.3. The Epidemiology of Cataract in Europe

The crude prevalence of cataract in European adults in 2007 was 19.3%. The presented prevalence rates highlight the impact cataract has on the population in Europe. Based on a previous review, the highest crude prevalence of cataract in adults was estimated to be in Germany (0.20 per 100,000 person-years) followed by Italy (0.065 per 100,000 person-years) and then the European North of Russia (0.039 per 100,000 person-years) (107). As for the UK, the North London Eye Study showed a steadily increasing prevalence of cataract in different age groups: 16% in the 65 to 69 year age group, 24% in people of 70 to 74 years of age, 42% in those 75 to 79 years of age, 59% in those 80 to 84 years and 71% in people aged 85 years or more. It was estimated that 225,000 new cases of visually impairing cataract should be expected each year, the 5-year cumulative incidence being estimated at 1.1 million new cases among the population aged 65 years and older (51).

1.9. Prevention, Treatment and Care Pathway of Cataract

Although many advances in recent years have been made to identify the risk factors for cataract, there are no medical treatments proven to reverse, or even slow down, the progress of cataract (51). A number of studies suggest certain nutrients and nutritional supplements may reduce the risk of cataracts. One large, 10-year study of female health professionals found that higher dietary intakes of vitamin E and the carotenoids lutein and zeaxanthin from food and supplements were associated with significantly decreased risks of cataract (97). Other studies have shown that antioxidant vitamins such as vitamin C and foods containing omega-3 fatty acids, may reduce cataract risk (93, 94). Another method to reduce your risk of cataracts is to wear protective sunglasses blocking the suns UV rays when you are outdoors (108). These possible preventive measures have not been conclusively demonstrated to work and may be of limited use in a well-nourished population from a country with a temperate climate.

Surgery seems to be the only effective treatment to restore or maintain vision for cataract patients. Phacoemulsification is the most common technique used in developed countries.

The internal lens is emulsified with an ultrasonic hand piece and aspirated from the eye, then an intraocular lens implant (IOL) is placed into the remaining lens capsule (109). Intracapsular cataract extraction (ICCE) is carried out to remove the entire lens including its capsule. The procedure has a relatively high rate of complications due to the large incision required and pressure placed on the vitreous body. In countries where operating microscopes are available, this technique has been superseded by extra capsular cataract extraction (ECCE) where the nucleus and cortex are removed through the anterior capsule, leaving the posterior capsule in place, which produces faster visual rehabilitation and fewer complications (110).

Every year approximately 10 million cataract operations are carried out in the world. In a Cataract Surgery Statistics study, the cataract surgery rate in India has approximately doubled in the last 10 years to around 3,000 operations per million populations per year now. In the most developing countries of Asia, the current cataract surgical rate is between 500 and 1,500, and in many countries of Africa the rate is less than 500. For Western Europe, the cataract surgery rate is 4,000, and in North America the cataract surgery rate remained at 5,500 (111).

Figure 3 provides a brief process of the care pathway for cataract patients within the National Health Service (NHS) (124). Cataract management is a multi-professional process involving ophthalmologists, optometrists, nurses and technicians. The ophthalmologist should take the responsibility for diagnosis and management of the patient. The decision on whether to proceed to surgery should be made by the patient in discussion with an ophthalmologist. Cataract surgery should be performed by an ophthalmic surgeon while much of the process may be undertaken by the non-medical members of the team provided that they are properly trained and supervised (11).

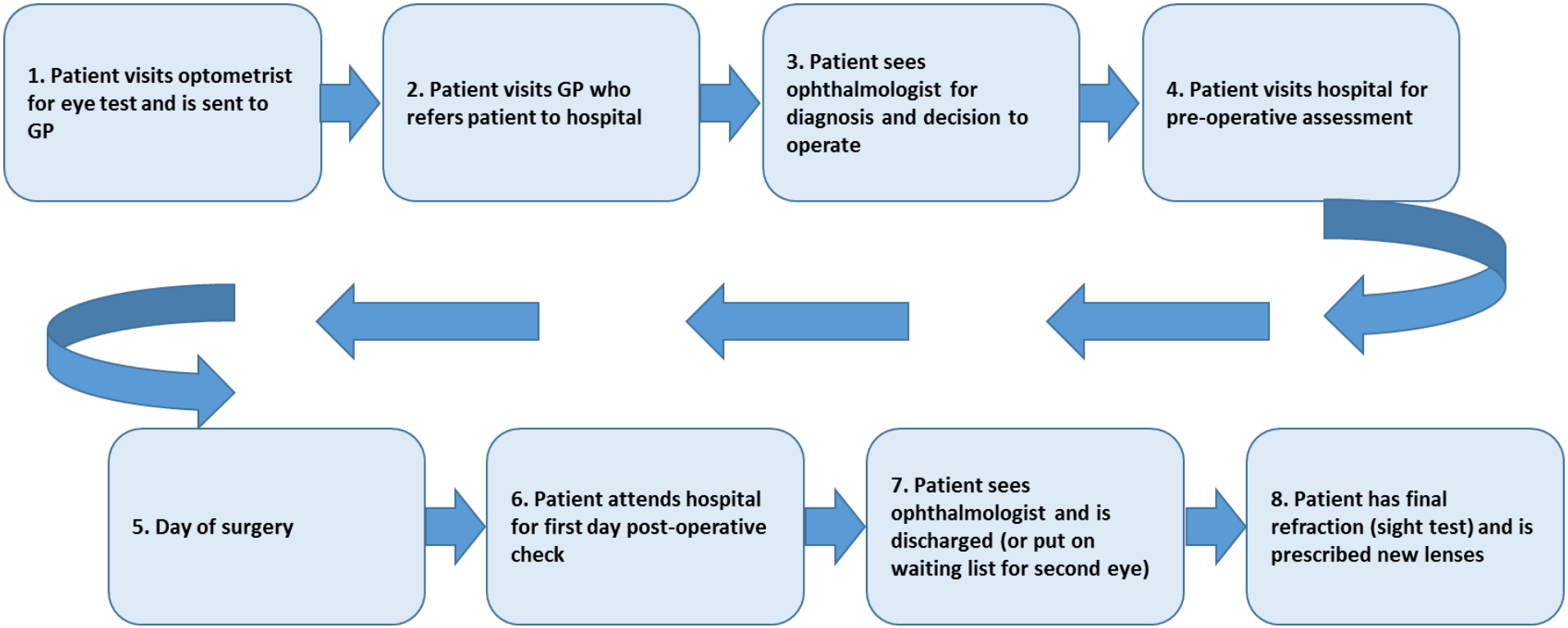

Figure 1.3: Process of the care pathway within the NHS

Chapter 2. Methodology

2.1. Study Setting and Design

2.1.1. Introduction of Genetics of Diabetes Audit and Research Tayside (GoDARTs) dataset

The Genetics of Diabetes Audit and Research Tayside (GoDARTS) project was originally created to record and analyze the risk factors for diabetes and its complications. This is an international biomedical and genetic resource for the study of type 2 diabetes and related conditions. It is estimated that there are now over 180 million people worldwide with diabetes. This chronic disease can cause serious health problems including heart disease, blindness, kidney failure and amputation. A high quality resource, initially funded by the Wellcome Trust and supported by Diabetes UK, has been created with successful recruitment of consenting patients with type 2 diabetes and matching controls (non-diabetics) throughout Tayside. This resource is already available to researchers worldwide and is helping to define genetic factors related to diabetes including susceptibility, complications and response to treatment. This international resource will become more powerful with time and will benefit from continual collection and the fast pace of technological progress in the field of genetic research. The response from the Tayside population has been exceptional and volunteers have participated in clinics, GP surgeries and work places throughout the region, and the project has now started recruiting in Fife. The participants consent to a baseline measurement-link to data gathered and anonymous follow up through datasets derived from medical records.

2.1.2. GoDARTS Data Collection

Every individual was required to complete a lifestyle questionnaire, a baseline clinical examination, and offer the biological samples when they came into the group. With their permission, the researchers would have access to their health information and biological samples for the study and a link to the private information the NHS medical records hold. A detailed history record, including demographics, prescribing history, general practice clinic visits, hospital admissions and outpatient appointments, was obtained at the base

hospital and was then linked with the Scottish Care Information-Diabetes Collaboration (SCI-DC) database; SCI-DC was commissioned and is owned by the Scottish Government. Every patient has a tracked electronic health record system and receives the care in Scotland.

2.2. Date Source and Participants

2.2.1. Case and Control Definition

The GoDARTS project has over 9,000 patients with type 2 diabetes and over 8,000 age-matched controls in the Tayside region of Scotland. For our project, we recruited 9,439 type 2 diabetic patients originally. Based on our project, the diabetic cataract case was defined as a type 2 diabetic patient recorded in the linked e-health records as having cataract in one eye at least or having previous cataract extraction surgeries in one eye at least. In fact, cataract appears more often in a mixed format mentioned above combining nuclear cataract, cortical cataract or posterior subcapsular cataract, than a single entity in clinical settings. It was reported that around one in three cataracts are a mixed type in a diabetic population (47). Therefore, the case used in our study was clearly described as any cataract. The control was defined as a type 2 diabetic patient who had never been diagnosed with cataract in the e-health records and had no history of cataract surgeries. We extracted the missing individuals and incomplete data. After filtering, there were 1,986 cases and 3,429 controls left for our project analysis. For the case and controls, individuals were chosen from the original database and we extracted the cataract individuals from all eye problems, and these cataract individuals were divided into people with cataract and people with cataract extraction. For the first group, they were cataract present, cataract absent, cataract present unknown eye and absent unknown. For the second extraction group, they were extraction and extraction eye unknown. Obviously, if we put these data into our regression model without filtration, it would be complex. Therefore, according to the case and control definition, after extraction and combination, there was a new variable; Diabetic cataract type containing two groups, which were diabetic cataract patient tagging as 1, and non-diabetic cataract patient tagging as 2.

2.2.2. GWAS Data Collection and Handling

The project GoDARTS adapted two types of DNA chips to genotype its diabetic individuals. The Affymetrix SNP6.0 chips were funded by the Wellcome Trust Case Control Consortium 2 (WTCCC2) project (112), and the Illumina OmniExpress chips were funded by the Surrogate markers for Micro- and Macro-vascular hard endpoints for Innovative diabetes Tools (SUMMIT) project (113). Genotype data quality controls were based on the standard protocols that were established for the WTCCC2 studies (112) and the SUMMIT studies (113).

2.2.3. Epidemiological Variables Collection and Handling

According to the literature review, important factors are required to find whether there is significant association between the potential risk factors and diabetic cataract, and to demonstrate an epidemiological result in the Tayside population. These variables for epidemiology analysis were also chosen from the original GoDARTs dataset.

Demographic Factors

Age: Age is the most common cause (114): lens proteins denature and degrade over time, and this process is accelerated by diabetes mellitus. In a previous study, the increasing age group is associated with a higher prevalence of cataract (98).

Gender: The combined evidence from the National Health And Nutrition Examination Survey (NHANES) follow-up study shows the excess risk for females to develop cataract (52). In a population-based prevalence survey in Beaver Dam, Wisconsin, women had more cortical opacities compared to men within similar age groups (115).

BMI: Body Mass Index (BMI) is a simple index of weight-for-height that is commonly used to classify underweight, overweight and obesity in adults. It is defined as the weight in kilograms divided by the square of the height in meters (kg/m2) (116).

Smoking status: From literature review, there are about 8 studies (73, 117, 118) that have

shown the association between increased risk of lens opacity and smoking.

Alcohol: UK government alcohol guidelines state that men and women are advised not to regularly drink more than 14 units per week. One unit equals 10ml or 8g of pure alcohol, which is around the amount of alcohol the average adult can process in an hour. This means that within an hour there should be, in theory, little or no alcohol left in the blood of an adult, although this will vary from person to person (119).

Economic status: We used the Scottish index of multiple deprivation (SIMD) to demonstrate the economic status of the studied population (120). It was a deprivation score based on the relative ranking within certain areas from most deprived to least deprived. In this study, Quintile SIMD includes five levels: “most deprived”, “deprived”, “middle”, “affluent” and “most affluent”.

Biochemistry Factors

We selected the total serum cholesterol (mmol/L), serum HDL cholesterol (mmol/L), serum LDL cholesterol (mmol/L), HbA1c m%, serum triglycerides (mmol/L) and blood pressure, which are divided into Diastole (mmHg) and Systole (mmHg), from biochemistry tests. In fact, we chose urine creatinine as one of the variables at the beginning, however there was more than 20% missing data that would influence the results; hence it was decided to exclude it.

2.3. Statistics Analysis

2.3.1. GWAS Statistics

The Linux software including Putty and WinSCP were used to deal with the original gene data. We collected the genotyped SNPs from Affymetrix and Illumina chips. Standard quality control steps such as exclusion of participants with more than 5% missing genotype data, SNPs with missing genotype of more than 5%, SNPs with less than 1% minor allele frequency and SNPs that failed Hardy–Weinberg tests (P<0.000001) were applied during data analysis using PLINK, which is the primary software for data

manipulation (121). SNPs on sex chromosomes and mitochondria were also excluded. *P* values were calculated by the logistic regression tests integrated in PLINK and *P* value of less than 10^{-6} was considered to be statistically significant. Based on the principle of GWAS, the *P* value is one of the key parameters to test the significance of association in the study. In fact, the exact threshold of *P* value varies by study, but the conventional threshold is 5×10^{-8} to be significant in the face of hundreds of thousands to millions of tested SNPs (121). Although the genome-wide significance *P* value threshold of 5×10^{-8} has become a standard for common-variant GWAS, it has not been updated to cope with the lower allele frequency, which means, in many GWAS studies where there is not enough population size to get more significant test value, we still are allowed to use wider P value instead (47). At the same time, the P value is not the only testing value in GWAS, we could also find the OR for each SNP in the studied population to find the significant association (45).

When obtaining the SNPs logistic results, we used the Haploview software to create the Manhattan plot, which displayed the 22 chromosomes on the X axis, and negative logarithm of the *P*-value for each SNP on the Y axis; each dot in the plot represents a SNP in this population. The strongest associations have the smallest *P*-values, and their negative logarithms will be the greatest. Additionally, we compared the frequency of related SNPs in our project with the 1000 Genome Project to recommend the dominant allele in each SNP.

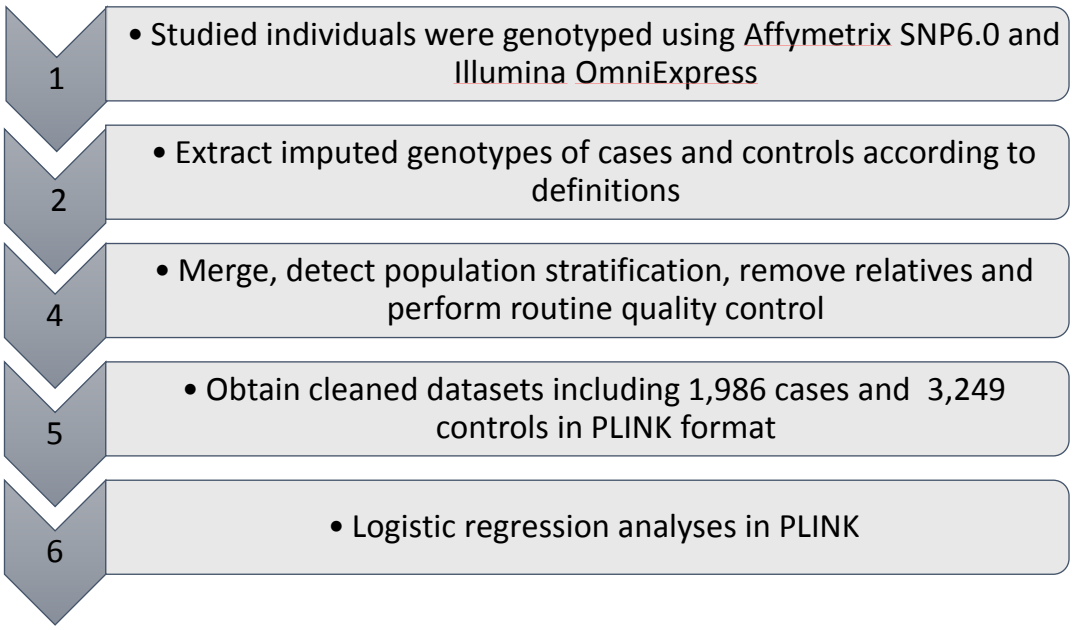

Figure 2.1: Workflow of the methodology for this GWAS on diabetic cataract using GoDARTS database.

2.3.2. Epidemiological Statistics

We put all variables into univariable and multivariable regression models to find the significant value for each variable. IBM SPSS software version 22 was used with regression models in analysis to find the association between potential risk factors and diabetic cataract. The X^2 test was used to test the difference of sex frequency between cases and controls and an independent t-test was used for other variables that were presented as mean ± standard deviation (Table 3.3).

There are six demographic variables including age, gender, BMI, smoking status, alcohol intake and SIMD. There are six biochemistry factors including Serum cholesterol, HDL-cholesterol, LDL-cholesterol, Blood pressure (diastolic and systolic), HbA1c and Triglycerides. All of these 12 variables were divided into continuous factors and categorical factors.

The categorical factors were classified as different levels and labeled in SPSS. For gender, 1 meant males and 2 meant females. For SIMD, 0 meant missing data, 1 represented the most affluent people, 2 represented affluent, 3 represented middle, 4 represented deprived, and 5 represented the most deprived people. For smoking status, 0 was unknown status,

1 was current smoker, 2 was ex-smoker, and 3 were people who had never smoked.

As for continuous variables, the normality test was used to check their normality respectively. For a better regression analysis, age, BMI, alcohol intake, total serum cholesterol, HDL-cholesterol, LDL-cholesterol, systolic pressure, diastolic pressure, HbA1c and serum triglycerides were transferred to a different variable with new classification. In a previous study (98), the variables were graded with quintiles; similarly, we classified these continuous variables with quartiles considering the sample size of our study, which could give us a clear classification and understanding about the relation between these factors and diabetic cataract. We used the 25%, 50% and 75% quartile to classify them into four groups; these new categorical variables were then put into our regression model. The alcohol intake was divided into two parts with 14 points because, based on the UK NHS guideline (119), men and women are advised not to regularly drink more than 14 units a week to reduce the risk of harming health (NHS live well). Therefore, 1 meant alcohol intake less than 14 units per week and 2 meant alcohol intake more than 14 units per week. We then put all variables into the univariable regression and multivariable regression models.

The univariable regression was used to test the association between the diabetic cataract and each potential risk factor, which is easy to understand. Moreover, we'd like to use multiple logistic regression when having one nominal and two or more measurement variables. The nominal variable is the dependent (diabetic cataract) variable; we are studying the effect that the independent (risk factors) variables have on the probability of obtaining a particular value of the dependent variable. The multiple logistic regression is to find an equation that best predicts the probability of a value of the Y variable as a function of the X variables we chose. Whether the purpose of a multiple logistic regression is prediction or understanding functional relationships, we usually want to decide which variables are important and which are unimportant, we can use an objective method (forward selection, backward elimination, or stepwise), or we can use a careful examination of the data and understanding of the biology to subjectively choose the best

variables. So we discussed the risk factors about their association in previous studies to decide which variable can be analyzed in the regression.

In the PLINK logistic regression, we used the Odds Ratio to describe the risk of potential SNP to the diabetic cataract in studied population. The GWAS also compare the association between SNPs and rare disease, which is more accurate than other parameters. Besides, in epidemiology statistics, logistic regression is one way to generalize the OR beyond two binary variables. And in our study, the OR is to quantify how strongly the presence or absence of the selected factor is associated with the presence or absence of the reference factor in a given population. The variables were classified as different levels and one level was chosen as reference. If the OR is greater than 1, the selected factor could rise the risk of developing the cataract compared with reference factor. On the contrary, if the OR is less than 1, the factor could reduce the risk of developing the cataract, which can be recognized as protective factor.

Otherwise, the relative risk (RR) is often used to express the association among the factors in cohort study. Both the OR and RR compare the likelihood of an event between two groups. However, our study is kind of case-control study, we wouldn’t choose the RR as the parameter, because, the RR measures events in a way that is interpretable and consistent with the way people really think. The RR, though, cannot always be computed in a research design. Also, the RR can sometimes lead to ambiguous and confusing situations.

Chapter 3. Results

3.1. Genetic Related Results

In this general diabetic cataract study of the Scottish population with GoDARTs, we identified 1,986 type 2 diabetic patients with cataract and 3,429 controls without cataract according to the linked e-health records after removing type 1 diabetic patients and patients with no genetic data. After filtrating, there were approximately 60,000 SNPs selected in this project to be analyzed by GWAS. Table 3.1 demonstrates all the significant SNPs from the smallest to largest in this population, whose significant value are less than 10^{-6}.

The most significant SNP is rs10197646 with *P* value of 4.12×10^{-7} (OR: 1.617, 95%CI: 1.343 1.948), which locates in the Chromosome 2. The next SNP is chr13: 48026216: D locating in Chromosome 13, this SNP has the second smallest *P* value of 4.15×10^{-7} (OR: 2.187, 95%: 1.615 2.961), but its OR is higher than other significant SNPs'. The rs62168795 (OR: 1.626, 95%CI: 1.342 1.947) and rs62168795 (OR: 1.416, 95%CI: 1.260 1.698) have a *P* value of 4.30×10^{-7} and 5.59×10^{-7}, respectively, which are all in the Chromosome 2. The *P* value of rs1381015 in Chromosome 4 dramatically increases to 7.12×10^{-7} when compared with the last SNPs, however the OR is the smallest among them (OR: 1.420, 95%CI: 1.236 1.630). Then, the rs2269547 in Chromosome 2 is close to the rs1381015 with the *P* value of 7.25x10-7 (OR: 1.809, 95%CI: 1.431 2.288). The last SNP is rs523355 (OR: 2.149, 95%CI: 1.585 2.915) with the largest *P* value of 8.63×10^{-7} and the largest OR among these 7 significant SNPs.

Table 3.1: Significant SNPs in the population

SNP	Chromosome	Odds Ratio	SE	Confidential Interval		*P*
				L95	U95	
rs10197646	2	1.617	0.0949	1.343	1.948	4.12×10^{-7}
chr13:48026216:D	13	2.187	0.1546	1.615	2.961	4.15×10^{-7}
rs7582173	2	1.616	0.0950	1.342	1.947	4.30×10^{-7}
rs62168795	2	1.463	0.0760	1.260	1.698	5.59×10^{-7}
rs1381015	4	1.420	0.0707	1.236	1.630	7.12×10^{-7}
rs2269547	22	1.809	0.1197	1.431	2.288	7.25×10^{-7}
rs523355	2	2.149	0.1555	1.585	2.915	8.63×10^{-7}

Table 3.2 shows the frequency of these 7 significant SNPs' dominant allele in case and control. All the SNP locate in the 2 allele. The frequency of dominant allele of rs10197646 is 0.1374 in case and 0.0940 in control (OR: 1.536). For chr13:48026216: D, the allele frequency in case (0.0548) is almost twice larger than that in control (0.0286), however, both of them are smallest among all alleles. The frequency of allele of rs7582173 in case is the same as that of rs10197646, but the frequency in control is slightly higher than that of rs10197646. As for rs62168795, its frequency in case and control are stable, being close to 20%. For rs1381015, both frequencies are more than 20%; moreover, the frequency in case is close to 30%. The SNP rs2269547 has a frequency of dominant allele of 0.1135 in case and 0.0668 in control. The last significant SNP has the highest OR, and the frequency in case and control are similar to the chr13:48026216:D.

The frequencies of all seven alleles (Table 3.2) in all case are higher than those in control, and the highest frequency in case and control for these alleles are all concentrating on the middle range. The frequency of each allele increases steadily, peaking at 0.2916 in case and 0.23300, then falls suddenly to the bottom. At the same time, the frequency of the most significant SNP's allele is slightly higher than the last significant SNP's allele, where there is no great difference between them.

Table 3.2: The Frequency of Allele in Case and Control

SNP	Chromosome	Dominant Allele	Frequency in Case	Frequency in Control	Odds Ratio
rs10197646	2	2	0.1374	0.0940	1.536
chr13:48026216:D	13	2	0.0548	0.0286	1.970
rs7582173	2	2	0.1374	0.0941	1.534
rs62168795	2	2	0.2355	0.1773	1.429
rs1381015	4	2	0.2916	0.2330	1.355
rs2269547	22	2	0.1135	0.0668	1.790
rs523355	2	2	0.0619	0.0310	2.060

The Manhattan plot of the GWAS for this project illustrates all associated SNPs genotyped in Affymetrix chips and Illumina chips from 1,986 cases and 3,429 controls in this project, which displays 22 chromosomes on X-axis with different colors, and negative logarithm of the P value for each SNP on Y-axis, where each dot in the plot represents a SNP in this population. As we know, the strongest associations have the smallest *P* values, and their negative logarithms will be the greatest. Since we set the significant value as 10^{-6} in this project, the seven significant SNPs mentioned in Figure 3.1 are above the red line and other insignificant ones are under it.

This plot (Figure 3.1) provides us a clear observation of the distribution of each SNP excluding ones with $P<0.01$ in the Manhattan plot of GWAS. There are seven SNPs reaching the GWAS significance ($P< 10^{-6}$). In the Chromosome 2, there are four SNPs, which are rs10197646, rs7582173, rs62168795 and rs523355; the smallest significant SNP and largest significant SNP are both located in this Chromosome. The rs10197646 and rs7582173 are too close to put together in the plot as they appear to be one dot. Other significant SNPs are located in Chromosome 4, 13 and 22, respectively.

As we can see, the majority of the *P* value of SNPs in the population are distributed from 10^{-2} to $10^{-4.5}$, although there are some separate dots in the range between 10^{-5} and 10^{-6} close to the significant line.

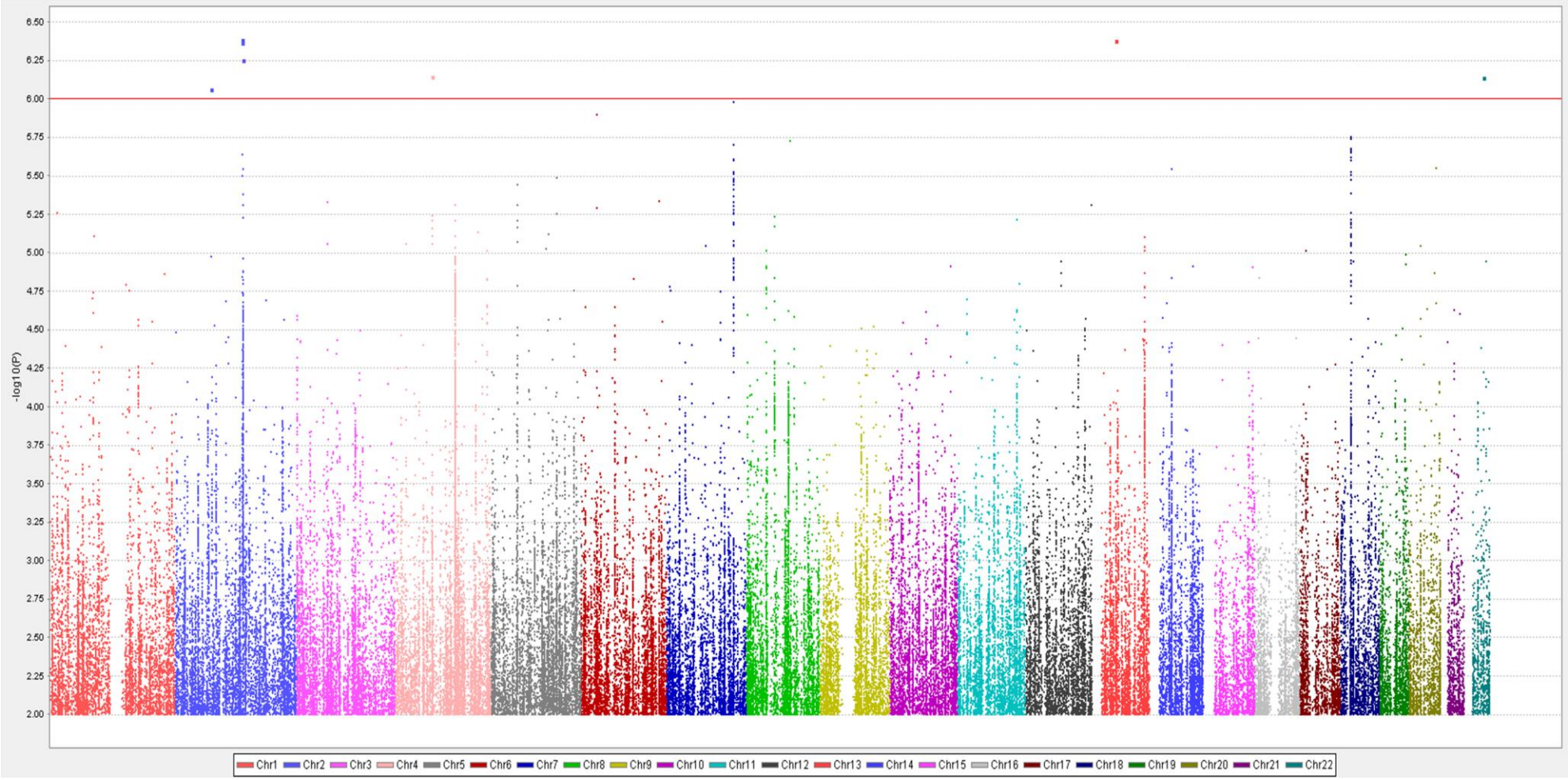

Figure 3.1: The Manhattan plot of the GWAS on diabetic cataract (1986 cases and 3429 controls) SNPs with $P < 0.01$ were excluded. The red line represents a *P value* of 10^{-6} in the plot.

3.2. Epidemiology Related Results

Table 3.3: Baseline Characteristics of Study Population

	Non-Cataract (3429)		Any Cataract (1986)		
	Mean± SD or N (%)	95% CI	Mean± SD or N (%)	95% CI	*P*
Age	66.45±9.82	66.12-66.78	70.01± 8.41	69.64-71.97	<0.001
Gender					0.0025
Male	1879 (54.8)	52.5-57.0	986 (49.6)	46.5-52.7	
Female	1549 (45.2)	42.7-47.6	1000 (50.4)	47.3-53.5	
BMI	31.27±5.50	31.0-31.4	30.64±4.70	30.4-32.6	0.002
Smoking					
Current smoker	314 (15.8)	11.7-19.8	346 (10.1)	6.9-13.3	
Ex-smoker	667 (33.6)	30.0-37.1	1348 (39.3)	36.7-41.9	
Never smoker	997(50.5)	47.4-53.6	1732 (50.5)	48.2-52.9	
Alcohol intake					
<14 per week	3407 (99.4)	99.1-99.6	1982 (99.8)	99.6-100	
⩾14 per week	22 (0.6)	0.1-3.8	4 (0.2)	0.1-4.6	
SIMD					
Most deprived	1013 (29.6)	26.7-32.4	442 (22.3)	18.4226.2	
Deprived	658 (19.2)	16.1-22.2	374 (28.3)	23.7-32.9	
Middle	446 (13)	9.8-16.1	303 (15.3)	11.3-19.4	
Affluent	724 (32.1)	28.7-35.5	516 (26)	22.2-29.8	
Most affluent	534 (15.6)	12.5-18.7	319 (16)	11.9-20.2	
Total serum Cholesterol, mmol/L	4.37±0.84	4.34-4.40	4.34±0.76	4.31-6.31	<0.001
Serum triglycerides, mmol/L	2.25±1.32	2.21-2.30	2.15±0.99	2.11-4.11	0.196
Serum HDL cholesterol, mmol/L	1.36±0.36	1.35-1.37	1.39±0.35	1.37-3.35	<0.001
Serum LDL cholesterol, mmol/L	2.05±0.65	2.03-2.08	2.03±0.59	2.01-3.99	0.04
HbA1c mg%	7.65±1.40	7.60-7.69	7.56±1.27	7.51-9.52	
Blood pressure					<0.001
Systole	136.92±29.88	135.92-137.92	139.23±30.83	137.87-140.58	<0.001
Diastole	73.10±20.34	72.42-73.79	71.68±20.66	70.77-73.64	0.002

Besides the genetic variants results, we also want to demonstrate some basic understanding of the epidemiological reports for the studied population to provide more additional demographic information in order to describe our population better.

Table 3.3 presents the baseline characteristics of the study population between non-cataract group (n=3,429) and any cataract group (n=1,986). Firstly, we should compare the difference of continuous variables between the non-cataract group and any cataract group. The average age of non-cataract people (66.45, 95%CI: 66.12-66.78) is younger than those with cataract (70.01, 95%CI: 69.64-71.97), however the standard deviation is higher than the cataract populations. Compared with the WHO BMI definition, the BMI results in our study show that the controls we chose had a slightly higher BMI average than cases, and its 95%CI range is smaller than that for cases. As for the biological test results containing total serum cholesterol, serum triglycerides, serum HDL cholesterol, serum LDL cholesterol, HbA1c and blood pressure, which are the keys factors in the health records in the GoDARTs, we might conclude that the average of these variables in the non-cataract population are all higher than those of any cataract population except for systolic blood pressure and serum HDL cholesterol.

The average of total serum cholesterol in controls is 0.03 bigger than that of cases (4.37 mmol/L vs. 4.34 mmol/L). The control groups serum triglycerides mean value is 0.1 higher than that of case group (2.25 mmol/L vs. 2.15 mmol/L). For serum HDL cholesterol and serum LDL cholesterol, the trend is opposite between the two groups. The HDL cholesterol in non-cataract population is lower than that in any cataract population (1.36 mmol/L vs. 1.39 mmol/L), but its standard deviation is larger in comparison (0.36 vs. 0.35). On the other hand, the average and standard deviation of LDL cholesterol in controls are both higher than those in cases. According to classification of blood pressure for adults ("Understanding blood pressure readings", American Heart Association), the desired blood pressure range is 90–119 mmHg for systolic blood pressure, and 60–79 mmHg for diastolic blood pressure. Back to our study, for the controls group, the blood pressure is 136.92 mmHg (Systole) (95%CI: 135.92 137.92) and 73.10 mmHg (Diastole) (95%CI: 72.42 73.79). Then any cataract people in our study hold the higher systolic blood pressure at 139.23 mmHg (95%CI: 137.87 140.58), and lower diastolic blood pressure at 71.68 mmHg (95%CI: 70.77 73.64). The categorical variables are shown with classification tables and charts below, which can provide more detail about the population

characteristics.

3.2.1. Age and Diabetic Cataract Prevalence

For better analysis, we transferred the age from the continuous variable to the categorical variable. According to Scotland's Census in 2011 from the Scottish government (122), the whole population was divided into multiple groups: 0-4 years old, 5-15 years old, 16-29 years old, 30-44 years old, 45-59 years old, 60-74 years old and 75 years old and over. However, in our studied population, the people were divided into four groups because the individuals are all over 30 years old (Figure 3.2).

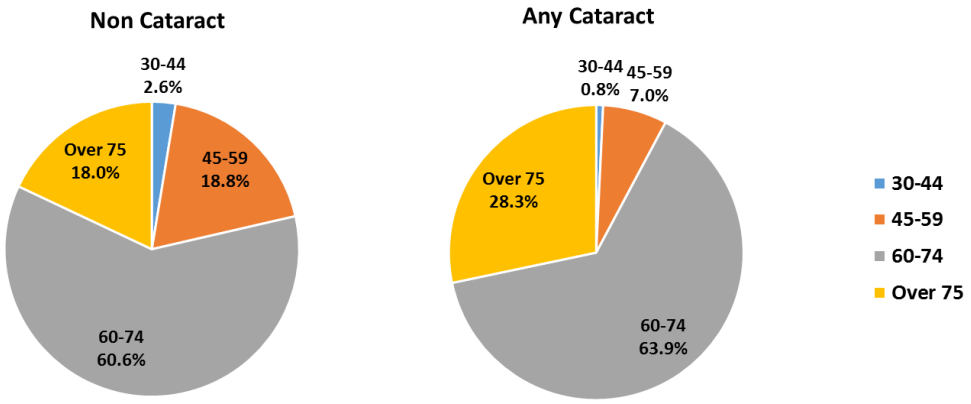

Figure 3.2: Percentage for each age group in the Non-cataract and Any Cataract groups

In the first group (30-44 years old), there are 89 patients without cataract and 15 cataract patients, where the occupation of controls (2.6%) is three times more than that of cases (0.8%), however both are the smallest in the Non-cataract group and Any cataract group. In the second group, between 45 and 59, the number of non-cataract patients rise to 645 occupying 18.8% of all controls. Meanwhile, the cases increase steadily and only take up 7.0% of cases population. In our studied population, most of the people are aged between 60 to 74 years old (almost 60% of the population in controls and cases groups (controls: 2,017, cases: 1,269)); although this age group is not the biggest part of the Scottish population census in 2011 (122). Then, for the last group, the number of non-cataract patients over 75 years old falls to a similar percentage (18.0) as the 45-59 group, and the cataract patients account for 28.3% of all cases in last age group.

3.2.2. Gender

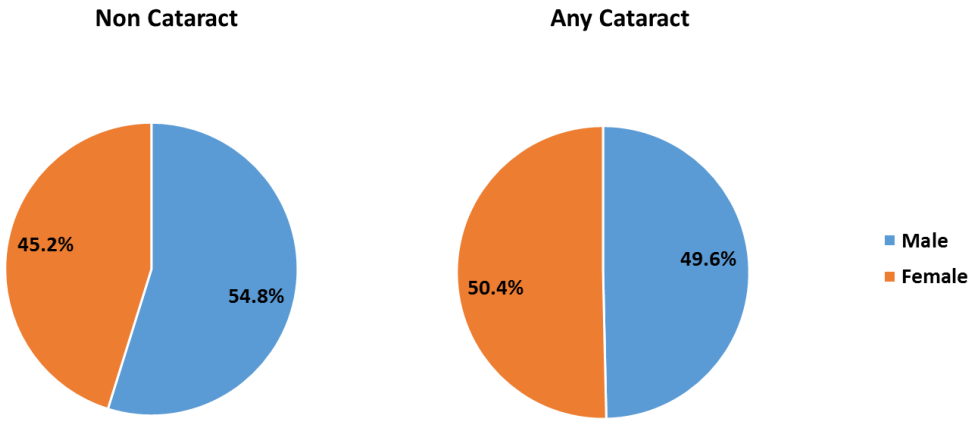

Figure 3.3: Percentage for gender in the Non-cataract and Any Cataract groups

When comparing the differences between males and females in our studied population, firstly we demonstrate the number and percentage of male and female in the control and case groups. In the control group, the percentage of males is almost 10% higher than that for females. On the contrary, the number of male and female patients in the cataract group is similar; the pie chart also shows that the percentage is almost 50% for each (Figure 3.3).

Table 3.4: Prevalence of diabetic cataract in male and female

Cataract	N	n	%	95% CI	RR
Male	2865	986	34.4	32.7-36.1	1
Female	2549	1000	39.2	37.3-41.1	1.14

Secondly, we found the prevalence of diabetic cataract between males and females. The list above (Table 3.4) shows the number of males (2,865) and females (2,549) in this study are similar. For the male group, the number of cataract patients is 986, accounting for 34.4% (95% CI: 32.7-36.1). The female cataract patients total 1,000 in the group with 39.2 % (37.3-41.1) cataract prevalence. Therefore, in this study, the cataract prevalence in females is a little higher than the prevelance in males, and the female group has the relative risk of 1.14 compared with the male group.

Table 3.5: Prevalence and age-adjusted prevalence in the Any Cataract and Non-cataract groups

		Prevalence		Age-Adjusted Prevalence	
Cataract	N	%	95%CI	%	95%CI
Any Cataract	1986	36.7%	34.6-38.8	24.9%	23.0-26.8
Non-Cataract	3429	63.3%	61.7-64.9	75.1%	73.7-76.5
Total	5415	100		100	

*Age adjusted to population of Tayside, based on census of Scotland Health Board Area 2011

The main epidemiological result in this project is demonstrated in Table 3.5 and shows the diabetic cataract prevalence is 36.7% in this population, and we used the Scotland health census age distribution to calculate the age-adjusted prevalence of 24.9%, which is lower (Table 3.5). The prevalence of each subtypes are not provided due to the lack of data in the initial health records.

3.2.3. Smoking Status and Alcohol

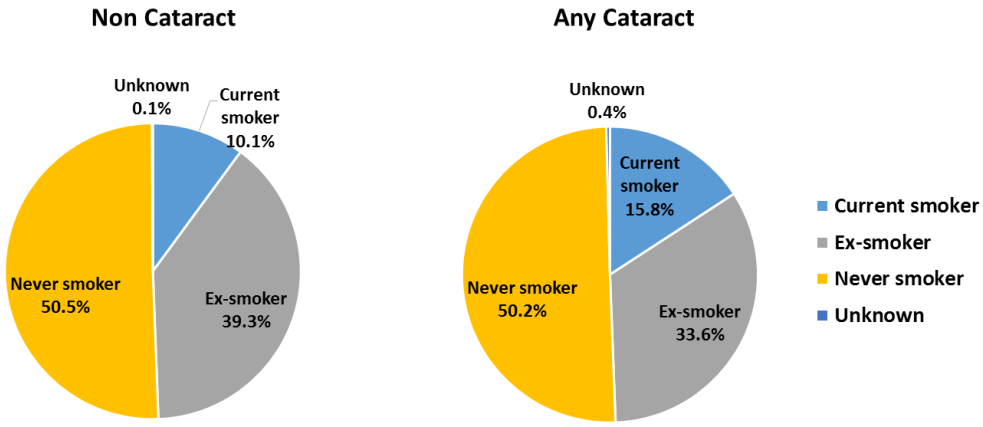

Figure 3.4: Percentage for each smoking status group in the Non-cataract and Any Cataract groups

The chart (Figure 3.4) above show the basic information regarding the smoking status of the population. As we can see, the group of never smokers is the main section in both the Non-cataract (50.51%) and any cataract groups (50.52%). The smallest section is for those of unknown status. For the current smokers, the people in the Non-cataract group

are 32 more than that in the any cataract group, but the percentage in Non-cataract people (10.1%) is higher than in the any cataract group (15.8%). For the Ex-smokers, the Non-cataract people total 1,348 accounting for 39.3%, which is the second largest section in the group. Meanwhile, any cataract people total 667 and occupy 33.9% in this group, which is half less than the amount of people in the non-cataract group, but slightly less than the percentage in it.

Table 3.6: Differences for alcohol intake in the Non-cataract and Any Cataract groups

	Non-Cataract		Any Cataract	
Alcohol	N	%	N	%
<14 per week	3407	99.4	1982	99.8
≥14 per week	22	0.6	4	0.2
Total	3429	100	1986	100

According the UK government alcohol intake guideline, the whole population was divided into two parts, and the vast majority of people have an alcohol intake lower than 14 unit per week in Non-cataract and Any cataract people (Table 3.6).

3.2.4. Scottish Index of Multiple Deprivation (SIMD)

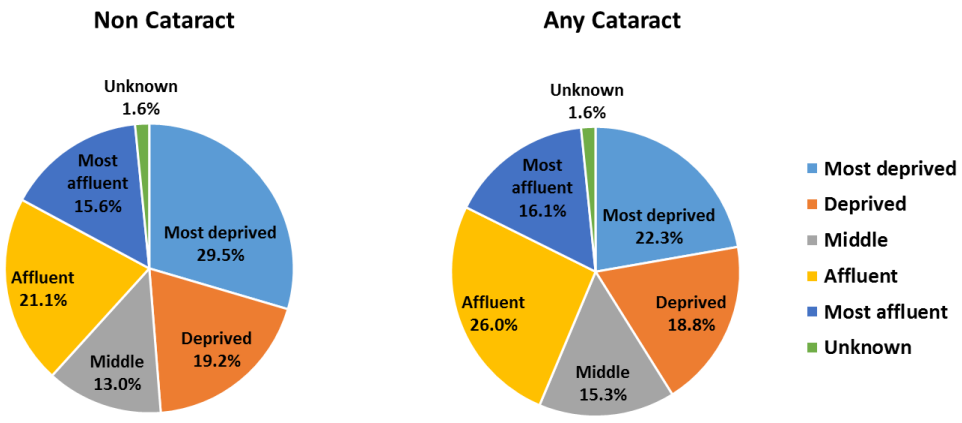

Figure 3.5: Percentage for SIMD in the Non-cataract and Any Cataract groups

We can see that the Scottish Index of Multiple Deprivation classification was divided into five levels. For these five groups, the number of people in the non-cataract group is higher than the any cataract group. However, in the non-cataract population, 29.5% people belong to the most deprived group, which is the largest section. Next, the deprived and affluent groups have a similar percentage of people (19.2% and 21.1%, respectively), then the most affluent and middle group only have 15.6% and 13.0% people, respectively. In the any cataract group, the largest section belongs to the affluent group with 26.0% cataract individuals, and the most deprived group still has the most cataract people (22.3%). The distribution in the deprived (18.8%) and most affluent groups (16.1%) are stable between non-cataract and any cataract group. Finally, the middle group is still the smallest group with 15.3%. Finally, we could ignore the unknown groups, which are only 1.6% in both cases and controls (Figure 3.5).

3.2.5. Biological Test

Table 3.7: Total serum cholesterol distribution in the Non-cataract and Any Cataract groups

	Non-Cataract		**Any Cataract**	
Total Serum Cholesterol mmol/L	N	%	N	%
1st Quartile (<3.92)	921	26.8	445	22.4
2nd Quartile (3.92-4.37)	730	21.3	643	32.4
3rd Quartile(4.38-4.65)	842	24.6	488	24.6
4th Quartile (>4.65)	936	27.3	410	20.6
Total	1986	100	3429	100

All continuous variables were classified into four groups with quantiles. For total serum cholesterol, the first (26.8%) and last groups (27.3%) have the most non-cataract people, on the contrary, in the cataract population, the middle groups (32.4% and 24.6%) account for the majority. Obviously, the distribution of non-cataract people increases with the rise of total serum cholesterol from second part peaking, but amount of any cataract people decreases with the rise of total serum cholesterol at the same point (Table 3.7).

Table 3.8: Serum HDL cholesterol distribution in the Non-cataract and Any Cataract groups

	Non-Cataract		Any Cataract	
Serum HDL Cholesterol mmol/L	N	%	N	%
1st Quartile (<1.16)	966	28.2	434	21.8
2nd Quartile (1.17-1.35)	1146	33.4	601	30.3
3rd Quartile (1.36-1.48)	450	13.1	494	24.9
4th Quartile (>1.48)	867	25.3	457	23.0
Total	1986	100	3429	100

If we look at the HDL cholesterol, the distribution of people is extremely imbalanced in the non-cataract population. There are 1,146 individuals (33.4%) in the second group but only 450 (13.1%) in the third group, and in the any cataract population, the distribution is similar in all groups except for the second (601, 30.3%). In surprise, the number of non-cataract people is much more than that of any cataract in the first, second, and fourth groups, but less than that in the third (Table 3.8).

Table 3.9: Serum LDL cholesterol distribution in the Non-cataract and Any Cataract groups

	Non-Cataract		Any Cataract	
Serum LDL Cholesterol mmol/L	N	%	N	%
1st Quartile (<1.72)	900	26.2	464	23.4
2nd Quartile (1.73-2.09)	1402	40.9	1007	50.7
3rd Quartile (2.10-2.22)	195	5.7	94	4.7
4th Quartile (>2.22)	932	27.2	421	21.2
Total	1986	100	3429	100

As for serum LDL cholesterol (Table 3.9, **错误!未找到引用源。**), the number of non-cataract (1,402) and any cataract people (1,007) peak at the highest point in the second group, moreover half percent of any cataract people are in this group. There are only 195 non-cataract individuals (5.7%) and 94 any cataract individuals (4.7%) in the third group,

which are the smallest of the whole population. For the other two groups, the numbers of non-cataract people (900, 932) are almost twice that of any cataract people (464, 421,Table 3.9).

Table 3.10: BMI distribution in the Non-cataract and Any Cataract groups

	Non-Cataract		Any Cataract	
BMI	N	%	N	%
1st Quartile(<27.70)	893	26.0	478	24.1
2nd Quartile(27.71-31.32)	824	24.0	691	34.8
3rd Quartile(31.33-32.70)	751	21.9	432	21.8
4th Quartile(>32.70)	961	28.0	385	19.4
Total	1986	100	3429	100

The BMI results show a stable distribution of studied people in all four BMI groups. In the Non-cataract group, the number of people decreases from 893 in the first group to 751 in the third group, but it eventually rises to 961 in the last group. For the any cataract group, the highest number (691) is in the second group and it dramatically falls down to 385 in the last group (Table 3.10).

Table 3.11: HbA1c distribution in the Non-cataract and Any Cataract groups

	Non-Cataract		Any Cataract	
HbA1c mg%	N	%	N	%
1st Quartile (<6.80)	976	28.5	493	24.8
2nd Quartile (6.81-7.50)	1099	32.1	893	45.0
3rd Quartile (7.51-8.00)	443	12.9	201	10.1
4th Quartile (>8.00)	911	26.6	399	20.1
Total	1986	100	3429	100

For HbA1c distribution, most of the people are located in the lower levels, 28.5% of controls in the first group and 32.1% of controls in the second group, while 24.8% of

cases in the first group and 45.0% of cases in the second group. The lowest number for both groups is in the third HbA1c group, 443, 12.9% and 201, 10.1%, respectively (Table 3.11, 错误!未找到引用源。).

Table 3.12: Triglycerides distribution in the Non-cataract and Any Cataract Groups

	Non-Cataract		**Any Cataract**	
Serum Triglycerides mmol/L	N	%	N	%
1st Quartile (<1.50)	893	26.0	479	24.1
2nd Quartile (1.51-2.23)	939	27.4	437	22.0
3rd Quartile (2.24-2.40)	633	18.5	688	34.6
4th Quartile (>2.41)	964	28.1	382	19.2
Total	1986	100	3429	100

In the serum triglycerides distribution table, the percentage of non-cataract population in all groups except the third group are similar (26.0%, 27.4%, 28.1%), the lowest point is 18.5% at the third group. For the any cataract population, the highest percentage is 34.6% at the third group, while in the other three groups, this trend decreases with the rise of the triglycerides level (Table 3.12).

Table 3.13: Blood Pressure shown for Systolic and Diastolic for the Non-cataract and Any Cataract Groups

	Non-Cataract		**Any Cataract**	
Systolic Blood Pressure mmHg	N	%	N	%
1st Quartile (<126)	901	26.4	450	22.7
2nd Quartile (127-140)	1032	30.1	569	28.7
3rd Quartile (141-152)	684	19.9	432	21.8
4th Quartile (>152)	809	23.6	535	26.9
Diastolic Blood Pressure mmHg				
1st Quartile (<68)	873	25.5	548	27.6
2nd Quartile (69-76)	838	24.4	528	22.6
3rd Quartile (77-83)	849	24.8	468	23.6
4th Quartile (>84)	869	25.3	442	22.3

Table 3.13 above illustrates the distribution of cases and controls with blood pressure, which is classified as systolic pressure and diastolic pressure (Table 3.13).

For the systolic pressure group, the largest percentage of controls is 30.1% in the second group, where the distribution of cases is also the largest (28.7%). The distribution in the other groups is similar. For the diastolic pressure group, the percentage in each group is similar for both cases and controls.

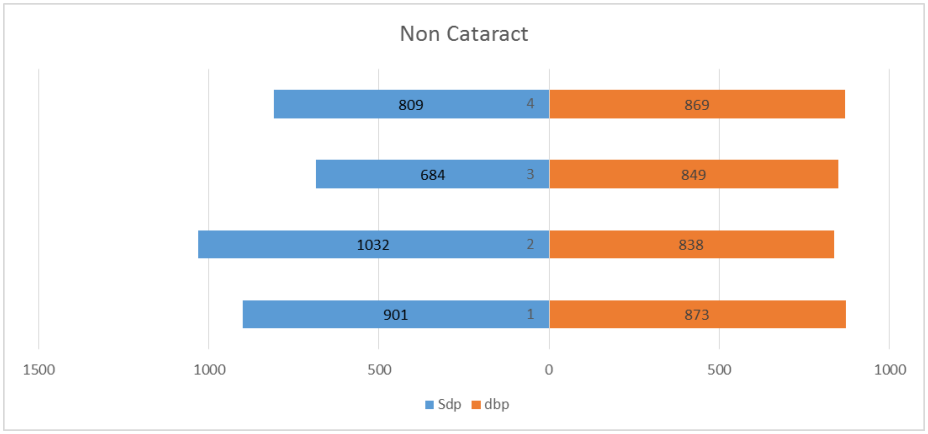

Figure 3.6: Comparison of systolic and diastolic pressure in the Non-cataract group

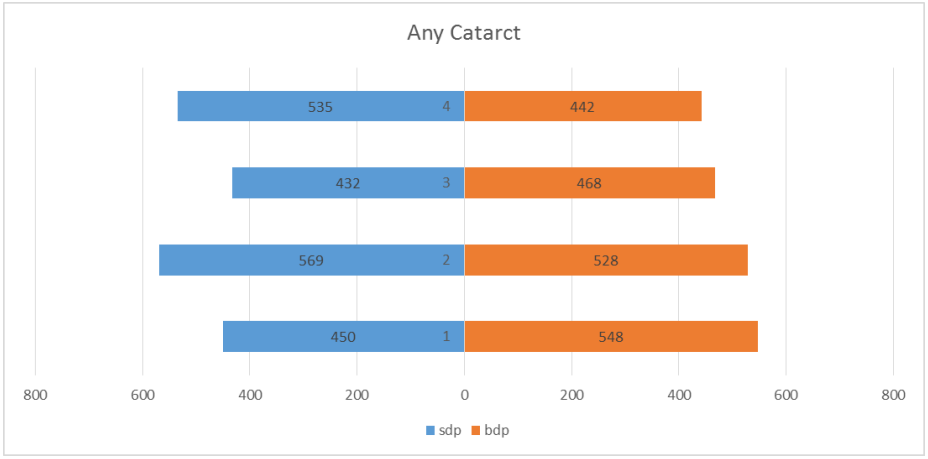

Figure 3.7: Comparison of systolic and diastolic pressure in the Any Cataract group

When comparing the systolic pressure and diastolic pressure together, it is clear that most of the controls have higher systolic pressure, and the people in diastolic pressure group is likely to be distributed equally, the amount of each part is around 800. At the same time, in the cases population, the number of cases in the diastolic pressure group goes up steadily from 442 to 548, while the number of cases in the systolic pressure group

fluctuates. In fact, the largest group is at the third part, and the smallest group belongs to the second part. In conclusion, the distribution of cases and controls in diastolic pressure are all more stable than that in systolic pressure (Figure 3.6, Figure 3.7).

3.2.6. Regression

Table 3.14 is the summery tables which include the results of the univariate analysis and multivariate analysis of all variables. The analysis was performed to identify risk factors and protective factors for diabetic cataract in our study population.

In the initial univariate analysis, the significant factors are age (P<0.001), gender (P=0.001), current smoker (P<0.001), alcohol intake (P=0.018), SIMD (P<0.001) and BMI (P<0.001). For biochemistry factors, the significant factors are pressure (Systole P=0.007, Diastole P=0.014), Total serum cholesterol (P<0.001), HDL serum cholesterol (P<0.001), LDL serum cholesterol (P<0.001), serum triglycerides (P<0.001) and HbA1c (P<0.001). The univariate analysis regression only shows the association with a single variable, where other cofounders and bias are excluded, which might affect the final significant association. Therefore, the results for every variable in the multivariate regression model could tell us better and give more understanding about the association between these factors and diabetic cataract.

In the multivariate regression table, age is significantly associated with diabetic cataract (P<0.001), and the odd ratio suggests that age is a protective factor to the diabetic cataract (OR: 0.955 95%CI: 0.948 0.962). In gender, being female is a possible risk factor related to the development of diabetic cataract (P=0.005, OR: 1.191, 95%CI: 1.055 1.345). In smoking status, current smokers are likely to have a higher risk of developing diabetic cataract (P=0.025, OR: 1.313 95%CI: 1.034 1.667), and non-smokers are also related to the diabetic cataract (P=0.048) as a risk factor. However, there is no evidence showing the relationship between ex-smokers and diabetic cataract.

Although SIMD is associated with diabetic cataract in univariate regression, the five

classifications with *P* value over 0.05 are not related to the diabetic cataract. Additionally, the most deprived group (OR: 1.171, 95%CI 0.733 1.871) and deprived group (OR: 1.010, 95%CI: 0.629 1.620) seem to have the risk for the diabetic cataract, but their 95% confidence interval with 1 cannot support the suggestion. Similarly, the odd ratio of affluent groups cannot prove the prevention for the development of diabetic cataract to some degree. Then, BMI factor is a significant factor for diabetic cataract. The first (P=0.183) and third (P=0.056) range are not significantly related to the cataract, but the second range (P=0.048, OR: 0.838, 95%CI: 0.703 0.998) is likely to be the protective factor for the progression of cataract.

When we come to the biochemical factors, total cholesterol serum, HDL serum cholesterol, Serum triglycerides and blood pressure are significantly associated with diabetic cataract in this regression model. In the total serum cholesterol groups, the lower serum cholesterol is a protective factor that might prevent the progression of diabetic cataract (2nd range, OR: 0.798, 95%CI: 0.642 0.992). The higher serum cholesterol ranges have no obvious relation to the cataract. In the HDL serum cholesterol groups, the higher range is strongly associated with diabetic cataract (3rd range, OR: 0.972, 95%CI: 0.723 1.307), which is also a protective factor for cataract; the higher serum triglycerides is strongly associated with the prevention of developing cataract in this population (P<0.001, OR: 0.393, 95%CI: 0.316 0.490). In blood pressure groups, the systolic blood pressure (P=0.003) and diastolic blood pressure (P=0.007) are both significant with diabetic cataract, moreover, systole is a protective factor in the study (OR: 0.997, 95%CI: 0.994 0.999), but diastole possibly increases the progression of diabetic cataract (OR: 1.004, 95%CI: 1.001 1.008).

Table 3.14: Risk factors for diabetic cataract in the regression analysis

	Univariate Analysis			Multivariate Analysis		
	OR	**95% CI**	***P***	**OR**	**95% CI**	***P***
Age	0.959	0.953-0.965	**<0.001**	0.955	0.948-0.962	**<0.001**
Gender Female	1.230	1.101-1.374	**0.005**	1.191	1.055-1.345	**0.005**
Blood Pressure						
Systole	0.997	0.996-0.999	**0.007**	0.997	0.994-0.999	**0.003**
Diastole	1.003	1.001-1.006	**0.014**	1.004	1.001-1.008	**0.007**
Smoking Status						
Current smoker	1.444	1.154-1.807	**0.001**	1.313	1.034-1.667	**0.025**
Ex-smoker	1.030	0.889-1.193	0.695	1.081	0.925-1.264	0.328
Never smoker	1.147	0.990-1.330	0.069	1.172	1.002-1.372	**0.048**
Unknown *(ref)*						
Alcohol Intake						
<14 per week	0.234	0.070-0.784	**0.018**	0.318	0.091-1.110	0.073
⩾14 per week *(ref)*						
BMI (kg/m²)						
1st Quartile <27.70	0.748	0.636-0.880	**<0.001**	1.133	0.943-1.361	0.183
2nd Quartile 27.71-31.32	0.478	0.409-0.558	**<0.001**	0.838	0.703-0.998	**0.048**
3rd Quartile 31.33-32.70	0.696	0.589-0.823	**<0.001**	1.238	0.995-1.541	0.056
4th Quartile >32.70 *(ref)*						
SIMD						
Most deprived	1.358	0.865-2.133	**<0.001**	1.171	0.733-1.871	0.510
Deprived	1.043	0.661-1.644	**<0.001**	1.010	0.629-1.620	0.969
Middle	0.872	0.550-1.383	**<0.001**	0.834	0.516-1.346	0.457
Affluent	0.831	0.529-1.306	**<0.001**	0.842	0.527-1.348	0.475
Most affluent	0.992	0.627-1.569	**<0.001**	0.995	0.617-1.603	0.982
Total Serum Cholesterol, mmol/L						
1st Quartile <3.92	0.907	0.771-1.066	**<0.001**	1.114	0.833-1.489	0.468
2nd Quartile 3.92-4.37	0.497	0.425-0.582	**<0.001**	0.798	0.642-0.992	**0.042**
3rd Quartile 4.38-4.65	0.756	0.643-0.888	**<0.001**	1.049	0.844-1.305	0.664
4th Quartile >4.65 *(ref)*						
Serum HDL Cholesterol, mmol/L						
1st Quartile <1.16	1.173	1.000-1.377	**<0.001**	0.897	0.728-1.107	0.311
2nd Quartile 1.17-1.35	1.005	0.865-1.168	**<0.001**	1.145	0.942-1.391	0.175
3rd Quartile 1.36-1.48	0.480	0.405-0.570	**<0.001**	0.737	0.596-0.910	**0.005**
4th Quartile >1.48 *(ref)*						
Serum LDL Cholesterol, mmol/L						
1st Quartile <1.72	0.876	0.746-1.029	**<0.001**	0.930	0.709-1.220	0.600
2nd Quartile 1.73-2.09	0.629	0.546-0.724	**<0.001**	0.856	0.700-1.048	0.132
3rd Quartile 2.10-2.22	0.937	0.714-1.230	**<0.001**	0.972	0.723-1.307	0.852
4th Quartile >2.22 *(ref)*						
HbA1c mg%						
1st Quartile <6.80	0.867	0.739-1.017	**<0.001**	1.035	0.872-1.229	0.692
2nd Quartile 6.81-7.50	0.539	0.465-0.624	**<0.001**	0.888	0.745-1.058	0.182
3rd Quartile 7.51-8.00	0.965	0.787-1.184	0.734	1.110	0.894-1.379	0.343
4th Quartile >8.00 *(ref)*						
Serum Triglycerides mmol/L						
1st Quartile <1.50	0.739	0.628-0.869	**<0.001**	0.814	0.656-1.012	0.063
2nd Quartile 1.51-2.23	0.851	0.723-1.003	0.055	0.938	0.775-1.135	0.508
3rd Quartile 2.24-2.40	0.365	0.311-0.428	**<0.001**	0.393	0.316-0.490	**<0.001**
4th Quartile >2.41 *(ref)*						

Significant results are shown in bold front (p<0.05).

Chapter 4. Discussion

4.1. Genetic Variants Discussion

In our studies, we performed a GWAS on diabetic cataract using a Scottish diabetic cohort based on phenotype information from linked e-health records and genetic information from DNA chips. All diabetic patients in Scotland attend retinal screening annually. During the screening, whether patients have cataracts or not will be determined by clinicians. However, when recording the case of a diagnosis of a cataract, they do not report the specific subtype of the cataract or the severity of the cataract. As we learned from background reading, the cataract appears more often in a mixed format (a combination of nuclear cataract, cortical cataract or posterior subcapsular cataract) rather than a single entity in clinics (123). Therefore, the phenotype used in our study was defined as ‘‘any cataract,’’ including mixed cataracts and any subtypes of cataracts.

In these results, we identified 7 significant SNPs associated with diabetic cataract.

The SNP of rs10197646 has the smallest *P* value of 4.12×10^{-7}, (OR 1.617, 95%CI 1.343, 1.948), located in the No. 2 Chromosome, with two kinds of allele, G/A; Allele A is the Ancestral one. (124). This SNP is greatly associated with diabetic cataract in this population, and the OR also shows the allele frequency in cases is much higher than that in the controls. In our study, the results demonstrated the basic understanding of the significant SNPs in our population. Therefore, the information of significant SNPs from the 1000 Genome Project Phase 3 was selected as the standardization to compare with the SNPs in our project. The 1000 Genomes Project was an international research with the goal of, firstly, creating a complete and detailed catalogue of human genetic variations, which in turn can be used for association studies relating genetic variation to disease, and secondly providing better support of SNP and probe selection for genotyping platforms in future studies and the improvement of the human reference sequence (125). The 1000 Genome Project compares the two alleles’ distribution. In the whole population, the A allele is 38.7%, and the G is 61.3%, where there is a clear difference between them. Then,

if we look at the European population, the distribution changes with the A allele with 25% and G with 75%. When we search the GBR population (British in England and Scotland) in the project, it is obvious that the difference is changing dramatically. The percentage of G allele is about four times more than that of A allele, (A 17%, G 83%). At this step, the distribution shows the distinct distribution of these two alleles adjusted to the different population. After allele frequency, the 1000 Genomes Projects also performs the genotype frequency as well, which are GG(39.3%), AA(16.7%) and GA(44.4%) in the whole population, GG(58.3%), AA(8.3%), and GA(33.4%) in the European population, then GG(70.3%), AA(4.4%), and GA(25.3%) in the British population. This suggests that the genotypes changing are related to the different race populations, and GG is the dominant genotype in the British (Table 4.1). All the results suggest that the G allele is the majority in British people. Although the results could not show the alleles' distribution and genotype directly in our studied population, we can still gather some basic information for the dominant allele for future studies. In our studied population, the special allele was located in the No.2 locus, and its frequency is 0.1374 in cases and 0.0940 in controls (Table 3.2).

Besides, in the same 2 chromosome, SNP of rs7582173 (P: 4.30×10^{-7} 95%CI: 1.342, 1.947) is closer to the rs10197646 SNP, which is suggested to contain the similar function with each other. As we found from the SNP database, the rs7582173 SNP includes Allele G/A as well, and with the same Ancestral A allele (126). The population genetics between standard 1000 Genomes Project and our studied project could tell us the allele frequency for G and A are the same for the whole population and sub-population (Table 4.1). When looking at our population (Table 3.2), we could see a slight difference, where the special allele has the same frequency of 0.1374 in cases, but a different frequency of 0.09405 in controls. Although compared to the general population, where there is no difference to the distribution for G/A between rs10197646 SNP and rs7582173 SNP, our project still shows us there is little change for these alleles in a particular group.

The SNP of rs62168795 is the next SNP with a smaller *P* value of 5.59×10^{-7} (OR 1.463,

95%CI 1.26 1.698) located in the No.2 chromosome, whose allele are T and C, and C as the Ancestral allele (127). In the Human Genomes Projects' database, two kinds of allele distribute as 16.1% for T allele and 84.9% for C allele. The distribution of T and C alleles are equal in the European population, but the frequencies change dramatically in the British population, namely T is 72% and C is 28%, which are opposite to the general population. When we look at the genotype frequency, the results demonstrate 7.9% of TT, 75.7% of CC, and 16.4% of TC in the whole population. In the European population, the frequency of the three genotypes is similar, TT 13.8%, CC 31.2% and TC 37%. The British people have the highest percentage with TT 53.8%, CT 36.3% and CC is only 9.9%. (Table 4.1) All the results show the T allele is dominant in the British population. As for the population we studied, the frequency of the dominant allele is 0.2355 in cases and 0.1773 in controls. The difference between them is smaller, which could mean this kind of allele is common in general people (Table 3.2).

The SNP of rs1381015 has $7.12x10^{-7}$ (OR 1.42, 95%CI 1.236 1.63) significant value in the population, which locates in the No.4 chromosome with A/C allele (Ancestral allele A) (128). Compared to the Human Genomes Projects' database, allele A is 87.3% and allele C is 12.7% for the whole population. For European people, allele A decreases to 75.6%, meanwhile, allele C increases to 24.4%. Similarly, these two types of alleles' distribution are stable in England and Scotland population (A 74.2%, C 25.8%). As for the genotype, allele A is the dominant allele in all genotype. The whole population genotype consists of 76.8% of AA, 21.1% of CC and 2.1% of AC (Table 4.1). The genotype in the European and British population are similar. In our population, the frequency of the dominant allele is closer in cases (0.2916) and in controls (0.233) (Table 3.2).

The rs2269547 SNP is associated with the particular disease in our project with $7.25x10^{-7}$ (OR 1.809, 95%CI 1.431 2.288). This SNP was defined in No.22 Chromosome with alleles G/C (Ancestral G) (129). In the 1000 Genomes Project Phase 3, the frequency of G (88.8%) allele is much higher than C (11.2%) in the whole population. We also found

that the distribution of the two kinds of alleles in the European and British population are similar to the whole population. Even the allele G has become 90.3% in England and Scotland, which suggests that most British people hold single G allele. This frequency also has an influence on the distribution of genotypes. According to the huge difference between allele G and C, it is obvious that there are only GG (81.3%) and GC (18.7%) genotypes in the British population. However, for the European population and whole population, the three kinds of genotype are similar to each other (Table 4.1). When looking at the population, the frequency of dominant alleles in cases (0.1135) is higher than that in controls (0.06675), which suggests the particular allele expresses more on the diabetic cataract patients (Table 3.2).

The last associated SNP is rs523355 locating in No. 2 Chromosome, with *P* value of $8.63x10^{-7}$ (OR 2.149, 95%CI 1.585 2.915), the odd ratio for this SNP is greatest among them, which suggest the highest frequency in cases. Compared with 1000 Genomes Project Phase 3, this SNP is related to allele G and C (Ancestral G) (130). In the general population, we could easily find the majority allele is G (90.8%), allele C has only 9.2 percent. At the same time, the genotype of the whole population is made up of GG (82.9%), CC, (1.2%) and GC (15.9%), which supports allele G being the dominant one. In the European population, allele G (93.1%) is still much higher than allele C (6.9%), and the CC genotype falls to 4%. When looking at British people, there are only GG (90.1%) and CG (9.9%) genotypes. We might conclude that allele G is dominant in English and Scottish population (Table 4.1).

The SNP Chr13:48026216:D seems to be discovered recently, because there is no information about it in the SNP database. Therefore we still need more researches to identify the population genetics about this SNP.

Table 4.1: Population genetics allele frequencies from 1000 Genomes Project Phase 3

Population	Allele: frequency		Genotype: frequency		
rs10197646					
All	G: 0.613	A: 0.387	G|G: 0.393	A|A: 0.167	A|G: 0.440
EUR	G: 0.750	A: 0.250	G|G: 0.583	A|A: 0.083	A|G: 0.334
GBR	G: 0.830	A: 0.170	G|G: 0.703	A|A: 0.044	A|G: 0.253
rs7582173					
All	G: 0.613	A: 0.387	G|G: 0.393	A|A: 0.167	A|G: 0.440
EUR	G: 0.750	A: 0.250	G|G: 0.583	A|A: 0.083	A|G: 0.334
GBR	G: 0.830	A: 0.170	G|G: 0.703	A|A: 0.044	A|G: 0.253
rs62168795					
All	T: 0.161	C: 0.839	T|T: 0.079	C|C: 0.757	C|T: 0.164
EUR	T: 0.503	C: 0.497	T|T: 0.318	C|C: 0.312	C|T: 0.370
GBR	T: 0.720	C: 0.280	T|T: 0.538	C|C: 0.099	C|T: 0.363
rs1381015					
All	A: 0.873	C: 0.127	A|A: 0.768	A|C: 0.211	C|C: 0.021
EUR	A: 0.756	C: 0.244	A|A: 0.573	A|C: 0.368	C|C: 0.060
GBR	A: 0.742	C: 0.258	A|A: 0.571	A|C: 0.341	C|C: 0.088
rs2269547					
All	G: 0.888	C: 0.112	G|G: 0.793	C|C: 0.018	C|G: 0.190
EUR	G: 0.883	C: 0.117	G|G: 0.779	C|C: 0.014	C|G: 0.207
GBR	G: 0.907	C: 0.093	G|G: 0.813	C|G: 0.187	
rs523355					
All	G: 0.908	C: 0.092	G|G: 0.829	C|C: 0.012	C|G: 0.159
EUR	G: 0.931	C: 0.069	G|G: 0.867	C|C: 0.004	C|G: 0.129
GBR	G: 0.951	C: 0.049	G|G: 0.901	C|G: 0.099	

All: whole population EUR: European GBR: British in England and Scotland

Based on the SNP database, the rs10197646 and rs7582173 are related to the *MAP3K19* gene (official name *mitogen-activated protein kinase kinase kinase 19)*, also known as *SPS1/STE20-Related Protein Kinase YSK4; YSK4 Sps1/Ste20-Related Kinase Homolog; RCK; YSK4* (131). The *MAP3K191* gene is a protein coding gene with related function of transferase activity, transferring phosphorus-containing groups and protein tyrosine kinase activity (131). The protein coded by this kind of gene originally belongs to the *Mitogen-activated protein kinase*, which are involved in directing cellular responses to a diverse array of stimuli, such as mitogens, osmotic stress and heat shock and proinflammatory cytokines. They regulate cell functions including proliferation, gene

expression, differentiation, mitosis, cell survival and apoptosis (132). This kind of function may influence the process of cataract formation.

Recent findings in genome-wide association studies have enabled us to identify thousands of genetic variants that are associated with human diseases. In a previous GWASdb SNP-Disease Associations providing comprehensive functional annotations for each genetic variant, including genomic mapping information, regulatory effects, amino acid substitutions, evolution, gene expression and disease associations to classify these GVs according to diseases using Disease-Ontology Lite and Human Phenotype Ontology (133). Researchers tested the database with the significant value between at least $5.0x10^{-8}$ and $1.0x10^{-5}$ (134) in a GWAS study; the results showed that there are many related association between *MAP3K19* gene and diseases that may influence the diabetic cataract (133).

Cataract was the first priority we wanted to look at, but there is no direct evidence to the association between cataract and *MAP3K19* (135). Then in the GWASdb SNP-Disease Associations dataset, there are 3,604 genes associated with diabetes mellitus including *MAP3K* gene family, which may influence the development of cataract (133, 136). Persons with diabetes mellitus have been found to be at an increased risk of developing cataracts when compared with non-diabetic persons (85). In the sugar cataract formation, the intracellular increase of sorbitol leads to osmotic changes resulting in hydropic lens fibers that may degenerate and form sugar cataract (137, 138). Therefore, in this Osmotic hypothesis sugar cataract formation, aldose reductase mediated accumulation of polyols results in lens swelling leading to the development of cataract ultimately (3). Furthermore, there are other studies showing that osmotic stress in the lens caused by the diabetes induces apoptosis lens cells (139) that leads to the cataract (140).

Moreover, hypertension is associated with the *MAP3K* gene set from the GWASdb SNP-Disease Associations dataset (133, 141). In a cataract risk factors study in black population, the evidence proved that the hypertension is associated with the increased risk

of the development of cataract (142). Hypertension is a long term medical condition in which the blood pressure in the arteries is persistently elevated (143). Hypertension is usually classified as either primary (essential) hypertension or secondary hypertension (144); about 90–95% of cases are primary due to nonspecific lifestyle and genetic factors (144, 145). However, in our study, the general Caucasian population was analyzed to see what may cause the genetic difference when compared with black population. In this GWAS study, *MAP3K19* gene was defined to be associated with eye problems including Myopia and Age Related Macular Degeneration (146).

Myopia, a condition of the eye where light focuses in front instead of on the retina (147), causes distant objects to be blurry while close objects appear normal. Severe myopia was proven to increase the risk of cataracts (148). In a prevalence of cataract with high myopia in an Indian population study, the significant association between high myopia and nuclear cataract was established (OR: 3.8, 95% CI 2.9 5.2, P <0.001) (149), which is similar with The Blue Mountains Eye Study's observation (OR: 3.3%; 95% CI 1.5 7.4, P <0.001) (150). It declared that nuclear cataract was strongly associated with axial myopia, and the density of the nuclear cataract was higher in subjects with myopia (149). Also, several population-based studies among adults of different ethnicities offered strong evidence to support there is an association between nuclear cataract and myopia (151, 152).

Age-related macular degeneration (AMD) may result in blurred or no vision in the center of the visual field (153), which shows no symptoms at the early stage, however some people experience a gradual worsening of vision that may affect one or both eyes over time (154). Macular degeneration typically occurs in the older population, and genetic factors play an important role in the development that was defined in the GWASdb SNP-Disease Associations study (133). The severity of AMD is divided into early, intermediate and late types (154). The late type is additionally divided into dry and wet forms, and it is conformed to be one of the most common reason of blindness after cataract (155). In the Health and Nutrition Examination Survey, cataracts and macular degeneration may

compromise visual function in older Americans, for persons aged 65 to 75 years old, the incidence of cataract and macular degeneration both rise 3 or 4 times, which suggests there is an interaction between cataract and macular degeneration in the elderly (54). On the other hand, a cross-sectional study showed that cataract surgery was associated with increased prevalence of late AMD adjusting for age, race and sex. Furthermore, having a severe cataract in the eye was also associated with a slightly higher prevalence of late AMD (156). All of these studies supported the association between cataract and age-macular degeneration.

According to the Gene data (130), rs523355 SNP is on the *CCT7* gene, whose official name is chaperonin containing TCP1 subunit 7, also known as *Nip7-1, TCP1ETA, NIP7-1, Ccth, CCTETA, CCTH*. All of these synonyms will make it easy to look for the new findings for this gene in other studies. This gene belongs to the Heat shock protein family coding gene, which encodes a molecular chaperone that is a member of the chaperonin containing *TCP1* complex (CCT), also known as the *TCP1* ring complex (TRiC). This consists of two identical stacked rings, each containing eight different proteins. Another molecular function is that this genes can assist the folding of proteins upon ATP hydrolysis including actin and tubulin (157, 158). According to the gene expression database, as one of the heat shock protein encoding genes, *CCT7* expresses highly in the lens of camera-type eye in the human (159). The eye of the human being is similar to an adult lamprey, which possesses numerous features that are very similar to those of the eyes of jawed vertebrates. The lamprey's camera-like eye has a lens, an iris and extra-ocular muscles lacking intra-ocular muscles. Its retina also has a structure very similar to that of the retinas of other vertebrates, with important function cells in three nuclear layers (160). Additionally, the southern hemisphere lamprey possesses five morphological classes of retinal photoreceptor and five classes of opsin, each of which are closely related to the opsins of jawed vertebrates. Comparing these similarities, the expression of this gene affects transparent eye structure, which is the key point in the progress of cataract formation (160). In another investigation to test the physiological function of αA-Crystallin in lens cells, αA knock-out mouse lens epithelial cells and human lens epithelial

cells that over express αA were exposure in the high level of UV light, after four-hour exposure, The growth rate of αA(−/−) mouse lens epithelial cells was reduced by 50% compared with cells with αA. As a member of the small heat shock protein (sHSP) family, molecular Chaperone αA-Crystallin was found to enhance lens epithelial cell growth and resistance to UVA stress(161) (162). At the same time, the loss of molecular Chaperone αA-Crystallin was shown in the human lens at older age, which is likely to cause development of cataract with age-independent increase (163). Therefore, the expression of *CCT7* gene molecular Chaperone αA-Crystallin in human lens will affect the development of cataract at some point. Besides, we found that the *CCT7* gene has expression on human blood, kidney function, and Proteinuria, which all have different influence on the development and progress of cataract (133).

Based on the Geneview (164), the rs62168795 SNP is related to the *R3H domain containing 1* gene, also called *R3H domain containing 1*, which encodes *R3H domain-containing protein 1* protein. This kind of protein contains function of poly (A) RNA binding, interacting non-covalently with a poly (A) RNA, but there is no direct connection between this molecular function and cataract. If we search more about *R3HDM1* gene expression in the human cells' from previous experiments, we find this gene might express on the function of kidney, cardiovascular system, adipose tissue and so on (165). Instead, we found evidence showing that this gene expresses in the epithelium of lens (166) in the other animals. The experiment proved that apoptosis of lens epithelial cell could occur under the various stress factors including oxidative stress, UV and other toxic agents (140). There is a great deal of evidence the supports the death of the lens epithelial cells via apoptosis can lead to cataractogenesis. Firstly, the death of lens epithelial cells will disturb the lifelong growth of the lens, leading to the thinness of cataract lenses (167, 168) and the decreased density of epithelial cells in the cataract lenses (169, 170). Then, the elimination of homeostatic epithelial cell control of the fiber cells often occur when lacking patches of lens epithelial cells, which caused the impairment of the integrity and transparency of these fiber cells (171). These processes eventually enhance the aggregation of crystalline (172) related to development of cataract (173, 174). As a matter

of fact, this gene expression is on the rat lens for now, however there is still a connection because the structure of the lens system is similar between human and rats. In future, we need more studies to find the evidence of the expression of *R3HDM1* gene in human lens

The last gene we described is *GGA1*, with official name of *Golgi-associated, gamma adaptin ear containing, ARF binding protein 1,* which is related to the rs2269547 SNP. This gene encodes a member of the Golgi-localized, gamma adaptin ear-containing, ARF-binding (GGA) protein family (175). Members of this family are ubiquitous coat proteins that regulate the trafficking of proteins between the trans-Golgi network and the lysosome. These proteins share an amino-terminal VHS domain which mediates sorting of the mannose 6-phosphate receptors at the trans-Golgi network (176). In an early study, the trans-Golgi network was related to calcium activity in the human system. The Golgi apparatus is an established calcium store (177, 178), and its role in sorting and processing secretory and membrane proteins is highly sensitive to changes in calcium concentration within the lumen. Recent experiments performed show that calcium gradients across the Golgi play a fundamental part in intracellular calcium signaling and homeostasis (178). Furthermore, the calcium channel is expressed and distributed in the epithelium and cortical fiber cells in the mouse lens (179). The different levels of calcium have been reported to induce progressive cortical cataract formation and are associated with decreased lens weight in ex vivo mouse lenses (179). Meanwhile, blocking of calcium exchange will cause the proliferation of human lens epithelial cells (180), and researchers have confirmed there was more than a 23-fold increase in total lens calcium with cataract (181). Therefore, all of these studies show there is a connection between our special gene and calcium activity in the human lens; we could conclude this gene's expression will influence the development of cataract formation to some degree. Besides, this gene also expresses on the lens of camera-type eye, the function of which has been discussed.

Apart from the genes mentioned above, there are still two significant SNPs, chr13:48026216:D and rs1381015, we could find the related genes form SNPs database. In the future, we need further studies to search wider SNP database to locate the related

genes and find out the potential association among them.

4.2. Epidemiology Discussion

Now we will examine the findings according to the epidemiological results to demonstrate the association between genetic variants and characteristics of population.

In our epidemiological study, the prevalence of diabetic cataract is 36.7%, and we used the age adjusted to the population of Tayside based on the census of Scotland Health Board Area 2011, to calculate the age-adjusted prevalence of cataract as 24.9% (Table 3.5). The Tayside prevalence of cataract is higher than the age-adjusted prevalence, which suggests that the age is a strong significant factor to the diabetic cataract, and in our studied population, the prevalence of cataract is increasing with the rise of age, but the prevalence in age group over 75 has fallen. The past study to estimate cataract prevalence for 24,409,978 people in the United States age 40 and older, shows the prevalence in different races all follow the basic tendency that increase straight to over 80 years old (182). Although the highest prevalence of cataract is in the group of 60-74 years old in our study, the percentage of cataract people still increases with the rise of age in the population (Figure 3.2). In the North London Eye Study's evidence, the prevalence of visually impairing cataract rose steadily with age: 16% in the 65 to 69 year age group, 24% in people of 70 to 74 years of age, 42% in those 75 to 79 years of age, 59% in those 80 to 84 years, and 71% in people of 85 years or more (51). And in our report, we could notice the result for age in Figure 3.2 and Table 3.14 are not consistent. We already discussed that early study confirmed the risk of diabetic cataract would increase with the rise of age (Table 3.2). However, the age in table 3.14 showed the different result, According to my statistical analysis, I used the Enter method to force all variables into the regression model to analyze without considering the contribution of each variables confounding others, which might mislead the real association between risk of diabetic cataract and each factor we chose.

Additionally, a follow-up study shows the estimate of cataract prevalence is likely to rise

with age in general, according to the Beaver Dam Eye Study (183), although we did not offer the prevalence of subtypes in our study like the Beaver Dam Eye Study did. Then, in the multivariate regression model, we found that the odd ratio for age seems to be less than 1 (0.955), which might explain the decline of cataract prevalence from the group 60-74 to the group over 75. In conclusion, growing age is the most common risk factor (114), lens proteins denature and degrade over time, and this process is accelerated by diseases such as diabetes mellitus (184).

Another key factor related to the age-specific cataract prevalence is the duration of diabetes, which was not analyzed in our study. As the significant risk factor for cataracts in diabetic patients, the longer duration of diabetes would indicate the higher incidence of cataract (90). A further 2 year follow-up in The Blue Mountains Eye Study showed that diabetic patients with longer than 10 years disease had higher Relative Risk (3.3) to develop the cataract in the future (185). The reason we did not include the duration of diabetic is that we excluded diabetes cases with incomplete cataract information during the data processing. In actual health records, patients with a shorter diabetes history were less likely to develop cataract, those who had no cataract tended to have an incomplete cataract record. As a result, the total number of shorter diabetes duration cases was smaller compared to cases for longer diabetes duration; many shorter diabetes duration cases without cataract were not included. Like the results in previous studies, stronger associations for the shorter period may be expected if the association is attributable to the effects of poor vision because the level of visual impairment present at the time a fracture would be much better reflected by a recent eye examination than one performed many years in the past. It is possible that frailer individuals, and those with more severe diabetes, were more likely to die before the end of follow-up, which could have attenuated this effect (185).

Gender is another key associated factor for diabetic cataract. The cataract prevalence of women was 39.2% when compared with 34.4% of men. In the any cataract group, the percentage of female is still higher than that of male. (Table 3.4) Similarly, the studies for

cataract prevalence in the United States and Australian population, state that women had the higher cataract prevalence than men (55, 59), and the relative risk of female compared to male (RR: 1.14) also suggests that women suffer more risk of developing cataract than men in our population. Moreover, the previous U.S. study with eHealth records that is similar to our data records, showed the prevalence of cataract by gender in different age groups (186) where the cataract prevalence of female is higher than the prevalence of male. Latest years' work have already proved that the females have a higher incidence of cataract than the males, and the prevalence as well (187, 188). In fact, our study offered similar cataract prevalence by gender in the age groups, although the prevalence for females is lower than that of male in group of 30-44, the possible reason might be that we excluded the patients under 30 years old which caused missing and incomplete data.

Smoking and alcohol are analyzed to have significant association with diabetic cataract separately. However, alcohol intake was proven to non-significant with diabetic cataract in multivariate regression. As a risk factor for diabetic cataract, current smoking cases (OR: 1.313 95%CI: 1.034 1.667) are more than that of controls as based on the tables shown (Table 3.14). A review of the evidence revealed that smoking increases the risk of a particular type of cataract – nuclear cataracts. (189, 190). Smokers who tend to consume 20 or more cigarettes per day, are at least twice as likely to develop a nuclear cataract compared to never smokers (74). A smoker's risk of developing cataract increases with the amount smoked; cataracts are more severe in heavy smokers than in light smokers (89, 191). Another important finding is that ex-smokers are more than current smokers for cases and controls groups in our study (Figure 3.4). A follow-up study lasting 12 years found current smokers of more than 15 cigarettes per day had a 42% increased risk of cataract extraction (OR:1.42, 95% CI, 1.28-1.58) compared with never smokers after adjustment for age and other potential risk factors. After more than 20 years since stopping smoking, these people had a 21% increased risk of cataract extraction (OR: 1.21, 95% CI, 1.06-1.39) compared with never smokers (192), which claimed smoking cessation significantly decreased the risk for cataract extraction with time ($P<0.001$). The higher the intensity of smoking, the longer it takes for the increased risk to decline, but

we did not collect the intensity of smoking for current smokers and ex-smokers, this could be discussed in any future study.

As for alcohol, the regression showed it is a non-significant factor for diabetic cataract (Table 3.14), because we use the UK government alcohol intake guideline to define the new variable, which might affect the association between alcohol and risk of diabetic cataract. Recent work by a team from Boston University found that heavy consumption of alcohol seriously increased the risk of having cataract surgery. Moderate consumption of alcohol seemed to correlate to reduced odds of surgery (193). The most important part is that an adverse effect of alcohol was stronger among smokers, people who smoked and drank heavily had an increased prevalence of nuclear cataract (194).

Our study demonstrated BMI is related to diabetic cataract, and the average BMI in our groups is 31.27 (non-cataract) and 30.64 (any cataract). According to the WHO BMI classification, the normal range is 18.50-24.99 (116), a person with BMI over 30 belongs to the obese group. Besides, if we look at the distribution of people in different BMI groups, we discover that 24.1% of cases in under the 27.70 group and 34.4% of cases are in the 27.70-31.32 group; which might suggest almost half of diabetic cataract patients are in the normal BMI range. The Framingham longitudinal studies using the Taylor and West lens grading system suggested that higher levels of average BMI and increasing BMI over time were risk factors for cortical opacity. Increasing BMI over time was also associated with posterior subcapsular opacity (195, 196). When looking at the different ranges of BMI, the results showed that only range (27.70-31.32) was significant with diabetic cataract, which is a protective factor as well (OR:0.838, 95%CI: 0.703 0.998) (Table 3.14), although other ranges are non-significant, their odd ratio are all over 1. A survey (101) on Chinese adults in Tanjong Pagar of Singapore showed lower BMI was associated with cortical cataract and any cataract.

Compared with relationship between risk of nuclear opacity and body mass index from the Shihpai Eye Study, we could easily find there is a similar odd ratio as ours. As a

protective factor in our study, the patients in range of 27.70-31.32 should have been less than other ranges', on the contrary, the percentage of any cataract in the lower range is higher than those in other ranges in our study. The possible reason could be that there is classification in our study. We should not miss that the previous studies use the LOCIII system to classify the cataract patients as nuclear, cortical and posterior subcapsular. The figure only showed the relationship between nuclear cataract and BMI, so if we put all subtypes into the model in future research, the results may be different. All in all, our study, combined with recent works, support that BMI is an independent risk factor for cataract (195).

We also included social economic factors in our regression model which appeared to be very interesting, SIMD score was used to measure deprivation level and it incorporated seven different aspects (employment; income; health, education, skills, and training; geographic access to services; crime and housing) of deprivation, combining them into a single index. SIMD score was based on ranking within a certain area (120). The SIMD scores we used were within the Tayside Health Board area and Scotland. The previous studies presented that more deprived populations have a higher risk of developing cataract (197). The logistic regression result showed that the affluent group (OR: 0.842) and most affluent group (OR: 0.995) in the Tayside area were protection factors for cataract (Table 3.14). Reports from the WHO Commission on Social Determinants of Health have emphasized the link between social and health inequalities, lower socioeconomic status has been shown to be associated with a higher risk of eye health (198, 199). The affluent people are able to have access to better healthcare to prevent cataract, even they can carry the cataract surgery at early state of disease, which will reduce the prevalence of cataract (199). While we should also focus on relatively rich groups with 95%CI including 1, this could still be a risk factor for diabetic cataract in this study. The possible explanation could be that rich people had higher chance of being outside for vacation and had higher VU light exposure. Another possible factor to be considered is that the Tayside area includes the sunniest places in Scotland, where the average is about 1,500 hours per year, which is more than average in any other places (maximum 1,300 hours per year) in

Scotland (200). Consequently, people in the sunniest places tend to protect themselves from being exposed in the strong UV light damage that is recognized as a potential risk factor of cataract. But we still need more studies to explain the reason for relatively affluent in Tayside area as a protect factor.

Next, we would like to talk about the biochemistry factors in the study. The univariate and multivariate analysis offered 5 identified variables as potential risk or protective factors for diabetic cataract (Table 3.14). There were no clear cut diagnose standard for Hypertension, thus the only cardiovascular related variable in our study was blood pressure (diastole and systole). Based on a prospective study of blood pressure and risk of cataract, there is no strong association between cataract and blood pressure, which is subject to confounding by multiple risk factors (201). Moreover, latest findings referred that cataract and systolic blood pressure were significantly associated, but not with diastolic BP (202). According to our results, the systolic blood pressure was a protective factor (OR: 0.997), which would decrease risk when systolic blood pressure rises. Oppositely, the diastolic blood pressure was shown to be the risk factor for diabetic cataract (OR: 1.004) (Table 3.14) . As a matter of fact, the average blood pressure (diastole and systole) in our population was higher than the desired blood pressure levels (Table 3.3), which is recognized as hypertension. Basically, hypertension could be recognized as the significant factor to cataract in previous studies (142). Under this circumstance, the cataract people seem to decrease with the rise of systolic blood pressure. As for diastolic blood pressure, it is a risk factor in our study (OR: 1.003), which would increase the risk of diabetic cataract.

As for three related variables, total serum cholesterol, serum HDL cholesterol and serum LDL cholesterol, our results showed that higher total serum cholesterol (2nd)seemed to decrease the risk of cataract (OR: 0.798, 95%CI: 0.642 0.992), then the serum HDL cholesterol is also a protective factor for diabetic cataract, because the 3rd quartile is strongly associated with diabetic cataract (OR: 0.737, 95%CI: 0.596 0.910), the LDL cholesterol is not significantly associated with the development of cataract in this study

though (Table 3.14). In many epidemiological studies, it has been suggested that, in different contexts, cholesterol may act as an antioxidant in the lens (203, 204). In our study, the higher total serum cholesterol could be considered as a protective factor. On the other hand, the high LDL cholesterol presents significant cataractogenic risk factors (89), but the lens is supported to be protective by the high serum HDL cholesterol (205), and the low LDL: HDL ratio also could decrease the risk of cataract in patients (205). However, we cannot find this possibility for HDL and LDL in our study, which might be due to the lack of whole study population's biological test.

The serum triglycerides are analyzed to have the possibility to prevent the development of diabetic cataract in higher level in our study. The distribution of cases in the higher serum triglycerides group also shows the same results that few cataract patients in higher serum triglycerides. The higher serum triglycerides could raise the risk of heart disease and stroke (206, 207) for the studied people, which would lead the death or incomplete health records. Unlike our study, the Framingham Study showed an association between elevated serum triglycerides and posterior cataract (98, 208). These differences in risk factors, among the different populations, can be related to the different genetic patterns associated with different types of cataract. We couldn't find accurate results about the association without the classification for the cataract in our study.

The limitation of the study was our inability to subtype the cataract groups firstly, which might invalidate some causal relationship between risk factor and cataracts, or miss the potential risk factors. Secondly, the wider population of Scotland was not taken into account. The SNPs were only collected in Tayside population from GoDARTs, which caused that we couldn't get more significant *P* value ($5x10^{-8}$) in our project. Then, our case definition includes all cataract patients with one or both eyes and patients with cataract extraction, we did not consider the prevalence of extraction and prevalence of cataract with one eye and that with both eyes, which could cause statistical error in our analysis. Although, most people will eventually develop a cataract in both eyes, one eye may be affected before the other. If taking this into consideration, we could predict that

most of the patients with cataract are still in the process of cataract development. Also, we used the ENTER to put all variables in to the logistic regression model which didn't consider the contribution of each factor. Over the past decades, rates of cataract surgeries have doubled in most of the UK; the current surgical rate approximates to a crude rate of 6.2 extractions per 1,000 population (11). The comparison with our results will show consistent condition for surgery rate, because the surgery may be the only effective treatment to cataract for now (11). All in all, further research is necessary to find more information about the diabetic cataract in the future.

Chapter 5. Conclusion

The key contribution of our study has two parts including genetic variants and epidemiological reports. In genetic variants contribution, we identified 7 significant SNPs for diabetic cataract in the Tayside area of the Scottish population with GoDARTs and found the related gene containing *MAP3K19*, *R3HDM1*, *GGA1*, and *CCT7* that associated with the progression of diabetic cataract. The epidemiological report provided the current cataract epidemiology status by excavating existing electrical health record. Major figures and trends are consistent with older studies in sources of population-based data for the prevalence of cataract in the UK, the prevalence of visually impairing cataract rose steadily with age (51, 209-212), and we also demonstrated other potential risk factors and protective factors of diabetic cataract in Tayside area in Scotland population.

With the established evidence in the fields of genetics, this project could provide more understandings about the mechanism of diabetic cataract, and drive us to carry out further related studies and functional studies to confirm the roles of more genes as risk factors associated with diabetic cataract. With the additional epidemiological reports, we should also focus more the preventing and treatment of diabetic cataract including details on the diagnostic criteria and guidelines for management in Scotland, and all of the UK. All in all, the findings give hopes for the millions of people with diabetes. They will not have to fear any possibility of life with visual impairment or blindness due to their illness and can live long, healthy, pleasurable, fulfilled lives.

References

1. Winslow C-E. The untilled fields of public health. Science. 1920:23-33.
2. WHO. Cataract: World Health Organization; 2016. Available from: http://www.who.int/topics/cataract/en/.
3. Pollreisz A, Schmidt-Erfurth U. Diabetic cataract—pathogenesis, epidemiology and treatment. Journal of ophthalmology. 2010;2010.
4. Porta M. A Dictionary of Epidemiology: New York: Oxford University Press; 2014.
5. Khoury MJB, Terri H.; Cohen, Bernice H. Fundamentals of Genetic Epidemiology: Oxford University Press; 1993 01-01.
6. WHO. Priority eye diseases: Cataract: World Health Organization; 2016. Available from: http://www.who.int/blindness/causes/priority/en/index1.html.
7. BlueEye. DEFINITION AND TYPOLOGIES: OPACITY OF THE CRYSTALLINE LENS, CONGENITAL CATARACTS AND AGE-RELATED CATARACTS. Available from: http://www.blueeye.it/en/eye-operations/cataract.php.
8. Sparrow J, Bron A, Brown N, Ayliffe W, Hill A. The Oxford clinical cataract classification and grading system. International ophthalmology. 1986;9(4):207-25.
9. Chylack LT, Wolfe JK, Singer DM, Leske MC, Bullimore MA, Bailey IL, et al. The lens opacities classification system III. Archives of ophthalmology. 1993;111(6):831-6.
10. Klein BE, Magli Y, Neider MW, Klein R. Wisconsin system for classification of cataracts from photographs: University of Wisconsin; 1989.
11. Ophthalmologists TRCo. Cataract Surgery Guidelines In: Ophthalmologists TRCo, editor. 2010.
12. Javadi M-A, Zarei-Ghanavati S. Cataracts in diabetic patients: a review article. Journal of ophthalmic & vision research. 2008;3(1):52-65.
13. Harding JJ, Egerton M, Van Heyningen R, Harding R. Diabetes, glaucoma, sex, and cataract: analysis of combined data from two case control studies. British Journal of Ophthalmology. 1993;77(1):2-6.
14. KAHN HA, LEIBOWITZ HM, GANLEY JP, KINI MM, COLTON T, NICKERSON RS, et al. The Framingham Eye Study II. Association of ophthalmic pathology with single variables previously measured in the Framingham Heart Study. American journal of epidemiology. 1977;106(1):33-41.
15. DE JAMA HPEP. Janet M. Torpy, MD, Writer Cassio Lynm, MA, Illustrator Richard M. Glass, MD, Editor.
16. Obrosova IG, Chung SS, Kador PF. Diabetic cataracts: mechanisms and management. Diabetes/metabolism research and reviews. 2010;26(3):172-80.
17. Dvornik D, Porte D, Ward JD. Aldose reductase inhibition: an approach to the prevention of diabetic complications: McGraw-Hill New York; 1987.
18. Varma S, Kinoshita J. The absence of cataracts in mice with congenital hyperglycemia. Experimental eye research. 1974;19(6):577-82.
19. Kador PF, Kinoshita JH. Role of aldose reductase in the development of diabetes-associated complications. The American journal of medicine. 1985;79(5):8-12.
20. Lee A, Chung SK, Chung S. Demonstration that polyol accumulation is responsible for diabetic cataract by the use of transgenic mice expressing the aldose reductase gene in the lens. Proceedings of the National Academy of Sciences. 1995;92(7):2780-4.
21. Kinoshita JH. Aldose reductase in the diabetic eye XLIII Edward Jackson Memorial Lecture. American journal of ophthalmology. 1986;102(6):685-92.
22. Burg MB, Kador PF. Sorbitol, osmoregulation, and the complications of diabetes. Journal of Clinical

Investigation. 1988;81(3):635.

23. Obrosova IG. Increased sorbitol pathway activity generates oxidative stress in tissue sites for diabetic complications. Antioxidants & redox signaling. 2005;7(11-12):1543-52.

24. Oates PJ. Aldose reductase, still a compelling target for diabetic neuropathy. Current drug targets. 2008;9(1):14-36.

25. Geisen K, Utz R, Grötsch H, Lang H, Nimmesgern H. Sorbitol-accumulating pyrimidine derivatives. Arzneimittel-Forschung. 1994;44(9):1032-43.

26. Kador PF, Inoue J, Secchi EF, Lizak MJ, Rodriguez L, Mori K, et al. Effect of sorbitol dehydrogenase inhibition on sugar cataract formation in galactose-fed and diabetic rats. Experimental eye research. 1998;67(2):203-8.

27. Sun W, Oates PJ, Coutcher JB, Gerhardinger C, Lorenzi M. A selective aldose reductase inhibitor of a new structural class prevents or reverses early retinal abnormalities in experimental diabetic retinopathy. Diabetes. 2006;55(10):2757-62.

28. Drel VR, Pacher P, Ali TK, Shin J, Julius U, El-Remessy AB, et al. Aldose reductase inhibitor fidarestat counteracts diabetes-associated cataract formation, retinal oxidative-nitrosative stress, glial activation, and apoptosis. International journal of molecular medicine. 2008;21(6):667-76.

29. Bron A, Sparrow J, Brown N, Harding J, Blakytny R. The lens in diabetes. EYE-LONDON-OPHTHALMOLOGICAL SOCIETY OF THE UNITED KINGDOM THEN ROYAL COLLEGE OF OPHTHALMOLOGISTS-. 1993;7:260-.

30. Lee S, Wang Y, Ko G, Ma R, Critchley J, Cockram C, et al. Risk factors for cataract in Chinese patients with type 2 diabetes: evidence for the influence of the aldose reductase gene. Clinical genetics. 2001;59(5):356-9.

31. Phillips SA, Mirrlees D, Thornalley PJ. Modification of the glyoxalase system in streptozotocin-induced diabetic rats: effect of the aldose reductase inhibitor Statil. Biochemical pharmacology. 1993;46(5):805-11.

32. Kim YS, Kim NH, Lee SW, Lee YM, Jang DS, Kim JS. Effect of protocatechualdehyde on receptor for advanced glycation end products and TGF-β1 expression in human lens epithelial cells cultured under diabetic conditions and on lens opacity in streptozotocin-diabetic rats. European journal of pharmacology. 2007;569(3):171-9.

33. Ranjan M, Nayak S, Rao BS. Immunochemical detection of glycated beta-and gamma-crystallins in lens and their circulating autoantibodies (IgG) in streptozocin induced diabetic rat. Mol Vis. 2006;12:1077-85.

34. Monnier VM, Stevens V, Cerami A. Nonenzymatic glycosylation, sulfhydryl oxidation, and aggregation of lens proteins in experimental sugar cataracts. The Journal of experimental medicine. 1979;150(5):1098-107.

35. Nakayama H, Mitsuhashi T, Kuwajima S, Aoki S, Kuroda Y, Itoh T, et al. Immunochemical detection of advanced glycation end products in lens crystallins from streptozocin-induced diabetic rat. Diabetes. 1993;42(2):345-50.

36. Mota M, Carvalho P, Ramalho J, Cardoso E, Gaspar A, Abreu G. Protein glycation and in vivo distribution of human lens fluorescence. International ophthalmology. 1994;18(4):187-93.

37. Franke S, Dawczynski J, Strobel J, Niwa T, Stahl P, Stein G. Increased levels of advanced glycation end products in human cataractous lenses. Journal of Cataract & Refractive Surgery. 2003;29(5):998-1004.

38. Gul A, Rahman MA, Salim A, Simjee SU. Advanced glycation end products in senile diabetic and nondiabetic patients with cataract. Journal of diabetes and its complications. 2009;23(5):343-8.

39. Biswas A, Miller A, Oya-Ito T, Santhoshkumar P, Bhat M, Nagaraj RH. Effect of site-directed

mutagenesis of methylglyoxal-modifiable arginine residues on the structure and chaperone function of human αA-crystallin. Biochemistry. 2006;45(14):4569-77.

40. Frank RN, Amin R, Kennedy A, Hohman TC. An aldose reductase inhibitor and aminoguanidine prevent vascular endothelial growth factor expression in rats with long-term galactosemia. Archives of Ophthalmology. 1997;115(8):1036-47.

41. Pop‐Busui R, Sima A, Stevens M. Diabetic neuropathy and oxidative stress. Diabetes/metabolism research and reviews. 2006;22(4):257-73.

42. Lee AY, Chung SS. Contributions of polyol pathway to oxidative stress in diabetic cataract. The FASEB Journal. 1999;13(1):23-30.

43. Obrosova IG, Drel VR, Oltman CL, Mashtalir N, Tibrewala J, Groves JT, et al. Role of nitrosative stress in early neuropathy and vascular dysfunction in streptozotocin-diabetic rats. American Journal of Physiology-Endocrinology and Metabolism. 2007;293(6):E1645-E55.

44. Obrosova IG, Mabley JG, Zsengellér Z, Charniauskaya T, Abatan OI, Groves JT, et al. Role for nitrosative stress in diabetic neuropathy: evidence from studies with a peroxynitrite decomposition catalyst. The FASEB journal. 2005;19(3):401-3.

45. Manolio TA. Genomewide association studies and assessment of the risk of disease. New England Journal of Medicine. 2010;363(2):166-76.

46. Pearson TA, Manolio TA. How to interpret a genome-wide association study. Jama. 2008;299(11):1335-44.

47. Chang C, Zhang K, Veluchamy A, Hébert HL, Looker HC, Colhoun HM, et al. A Genome-Wide Association Study Provides New Evidence That CACNA1C Gene is Associated With Diabetic CataractGenome-Wide Association Study and Diabetic Cataract. Investigative ophthalmology & visual science. 2016;57(4):2246-50.

48. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nature reviews genetics. 2008;9(5):356-69.

49. Groves CJ, Zeggini E, Minton J, Frayling TM, Weedon MN, Rayner NW, et al. Association analysis of 6,736 UK subjects provides replication and confirms TCF7L2 as a type 2 diabetes susceptibility gene with a substantial effect on individual risk. Diabetes. 2006;55(9):2640-4.

50. Delcourt C, Cristol J-P, Tessier F, Léger CL, Michel F, Papoz L, et al. Risk factors for cortical, nuclear, and posterior subcapsular cataracts: the POLA study. American journal of epidemiology. 2000;151(5):497-504.

51. Reidy A, Minassian D, Vafidis G, Joseph J, Farrow S, Wu J, et al. Prevalence of serious eye disease and visual impairment in a north London population: population based, cross sectional study. Bmj. 1998;316(7145):1643-6.

52. HILLER R, SPERDUTO RD, EDERER F. Epidemiologic associations with nuclear, cortical, and posterior subcapsular cataracts. American journal of epidemiology. 1986;124(6):916-25.

53. Hyman L. Epidemiology of eye disease in the elderly. Eye. 1987;1(Pt 2):330-41.

54. Klein BE, Klein R. Cataracts and macular degeneration in older Americans. Archives of Ophthalmology. 1982;100(4):571-3.

55. Group EDPR. Prevalence of cataract and pseudophakia/aphakia among adults in the United States. Archives of Ophthalmology. 2004;122(4):487.

56. Lewallen S, Mousa A, Bassett K, Courtright P. Cataract surgical coverage remains lower in women. British Journal of Ophthalmology. 2009;93(3):295-8.

57. Nirmalan PK, Krishnadas R, Ramakrishnan R, Thulasiraj RD, Katz J, Tielsch JM, et al. Lens opacities in a rural population of southern India: the Aravind Comprehensive Eye Study. Investigative ophthalmology

& visual science. 2003;44(11):4639-43.

58. Xu L, Cui T, Zhang S, Sun B, Zheng Y, Hu A, et al. Prevalence and risk factors of lens opacities in urban and rural Chinese in Beijing. Ophthalmology. 2006;113(5):747-55.

59. Mitchell P, Cumming RG, Attebo K, Panchapakesan J. Prevalence of cataract in Australia: the Blue Mountains eye study. Ophthalmology. 1997;104(4):581-8.

60. Leske MC, Chylack LT, Wu S-Y. The lens opacities case-control study: risk factors for cataract. Archives of Ophthalmology. 1991;109(2):244-51.

61. Leske MC, Chylack LT, He Q, Wu S-Y, Schoenfeld E, Friend J, et al. Risk factors for nuclear opalescence in a longitudinal study. American journal of epidemiology. 1998;147(1):36-41.

62. Chua J, Koh JY, Tan AG, Zhao W, Lamoureux E, Mitchell P, et al. Ancestry, Socioeconomic Status, and Age-Related Cataract in Asians: The Singapore Epidemiology of Eye Diseases Study. Ophthalmology. 2015;122(11):2169-78.

63. Sommer A, Tielsch JM, Katz J, Quigley HA, Gottsch JD, Javitt JC, et al. Racial differences in the cause-specific prevalence of blindness in east Baltimore. New England Journal of Medicine. 1991;325(20):1412-7.

64. Wesolosky JD, Rudnisky CJ. Relationship between cataract severity and socioeconomic status. Canadian Journal of Ophthalmology/Journal Canadien d'Ophtalmologie. 2013;48(6):471-7.

65. Echebiri S, Odeigah P, Myers S. Case-control studies and risk factors for cataract in two population studies in Nigeria. Middle East African journal of ophthalmology. 2010;17(4):303.

66. Kador PF, Kinoshita JH, editors. Diabetic and galactosaemic cataracts. Ciba Foundation Symposium 106-Human Cataract Formation; 1984: Wiley Online Library.

67. Harding JJ, Harding RS, Egerton M. Risk factors for cataract in Oxfordshire: diabetes, peripheral neuropathy, myopia, glaucoma and diarrhoea. Acta Ophthalmologica. 1989;67(5):510-7.

68. Chen T, Hockwin O, Dobbs R, Knowles W, Eckerskorn U. Cataract and health status: a case-control study. Ophthalmic research. 1988;20(1):1-9.

69. Hiller R, Kahn HA. Senile cataract extraction and diabetes. British Journal of Ophthalmology. 1976;60(4):283-6.

70. Ederer F, Hiller R, Taylor HR. Senile lens changes and diabetes in two population studies. American Journal of Ophthalmology. 1981;91(3):381-95.

71. DeBlack SS. Cigarette smoking as a risk factor for cataract and age-related macular degeneration: a review of the literature. Optometry (St Louis, Mo). 2003;74(2):99-110.

72. Bateman J, Rossi H, Kellerer AM, Robinson C, Bond V. Dose-dependence of fast neutron RBE for lens opacification in mice. Radiation research. 1972;51(2):381-90.

73. Christen WG, Manson JE, Seddon JM, Glynn RJ, Buring JE, Rosner B, et al. A prospective study of cigarette smoking and risk of cataract in men. Jama. 1992;268(8):989-93.

74. Kelly SP, Thornton J, Edwards R, Sahu A, Harrison R. Smoking and cataract: review of causal association. Journal of Cataract & Refractive Surgery. 2005;31(12):2395-404.

75. Ross M, Crosley L, Brown K, Duthie S, Collins A, Arthur J, et al. Plasma concentrations of carotenoids and antioxidant vitamins in Scottish males: influences of smoking. European journal of clinical nutrition. 1995;49(11):861-5.

76. Ramakrishnan S, Sulochana K, Selvaraj T, Rahim AA, Lakshmi M, Arunagiri K. Smoking of beedies and cataract: cadmium and vitamin C in the lens and blood. British journal of ophthalmology. 1995;79(3):202-6.

77. Cekic O. Effect of cigarette smoking on copper, lead, and cadmium accumulation in human lens. British Journal of Ophthalmology. 1998;82(2):186-8.

78. Mainster MA. Violet and blue light blocking intraocular lenses: photoprotection versus photoreception. British journal of ophthalmology. 2006;90(6):784-92.

79. Pitts DG, Cullen AP, Hacher P. Ocular ultraviolet effects from 295 nm to 400 nm in the rabbit eye. DHEW (NIOSH) Publication. 77: NIOSH; 1977.

80. Hollows F, Moran D. Cataract-the ultraviolet risk factor. The Lancet. 1981;318(8258):1249-50.

81. BRILLIANT LB, GRASSET NC, POKHREL RP, KOLSTAD A, LEPKOWSKI JM, BRILLIANT GE, et al. Associations among cataract prevalence, sunlight hours, and altitude in the Himalayas. American journal of epidemiology. 1983;118(2):250-64.

82. HILLER R, Sperduto RD, EDERER F. Epidemiologic associations with cataract in the 1971–1972 National Health and Nutrition Examination Survey. American journal of epidemiology. 1983;118(2):239-49.

83. Mohan M, Sperduto RD, Angra SK, et al. INdia-us case-control study of age-related cataracts. Archives of Ophthalmology. 1989;107(5):670-6.

84. Clayton R, Cuthbert J, Phillips C, Bartholomew R, Stokoe N, Ffytche T, et al. Analysis of individual cataract patients and their lenses: a progress report. Experimental eye research. 1980;31(5):553-66.

85. Klein BE, Klein R, Moss SE. Prevalence of cataracts in a population-based study of persons with diabetes mellitus. Ophthalmology. 1985;92(9):1191-6.

86. Rodríguez-Sargent C, Berrios G, Irrizarry JE, Estapé ES, Cangiano JL, Martínez-Maldonado M. Prevention and reversal of cataracts in genetically hypertensive rats through sodium restriction. Investigative ophthalmology & visual science. 1989;30(11):2356-60.

87. Cotlier E, Obara Y, Toftness B. Cholesterol and phospholipids in protein fractions of human lens and senile cataract. Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism. 1978;530(2):267-78.

88. Cenedella RJ. Source of cholesterol for the ocular lens, studied with U18666A: a cataract-producing inhibitor of lipid metabolism. Experimental eye research. 1983;37(1):33-43.

89. Heydari B, Kazemi T, Zarban A, Ghahramani S. Correlation of cataract with serum lipids, glucose and antioxidant activities a case-control study. West Indian Medical Journal. 2012;61(3):230-4.

90. Kim SI, Kim SJ. Prevalence and risk factors for cataracts in persons with type 2 diabetes mellitus. Korean Journal of Ophthalmology. 2006;20(4):201-4.

91. Cenedella RJ. Cholesterol and cataracts. Survey of ophthalmology. 1996;40(4):320-37.

92. Wilkins WKHLW. Cholesterol-lowering drugs may be linked to increased cataract risk. ScienceDaily. 2012.

93. Mares-Perlman JA, Brady WE, Klein BE, Klein R, Haus GJ, Palta M, et al. Diet and nuclear lens opacities. American journal of epidemiology. 1995;141(4):322-34.

94. Cumming RG, Mitchell P, Smith W. Diet and cataract: the blue mountains eye study. Ophthalmology. 2000;107(3):450-6.

95. Vitale S, West S, Hallfrisch J, Alston C, Wang F, Moorman C, et al. Plasma antioxidants and risk of cortical and nuclear cataract. Epidemiology. 1993:195-203.

96. Knekt P, Heliövaara M, Rissanen A, Aromaa A, Aaran R-K. Serum antioxidant vitamins and risk of cataract. Bmj. 1992;305(6866):1392-4.

97. Chasan-Taber L, Willett WC, Seddon JM, Stampfer MJ, Rosner B, Colditz GA, et al. A prospective study of carotenoid and vitamin A intakes and risk of cataract extraction in US women. The American journal of clinical nutrition. 1999;70(4):509-16.

98. Raman R, Pal SS, Adams JSK, Rani PK, Vaitheeswaran K, Sharma T. Prevalence and risk factors for cataract in diabetes: Sankara Nethralaya Diabetic Retinopathy Epidemiology and Molecular Genetics Study, report no. 17. Investigative ophthalmology & visual science. 2010;51(12):6253-61.

99. Clare Gilbert PA, Serge Resnikoff, Suzanne Gilbert,, Jill Keeffe CC, Ramachandra Pararajasegaram,

Silvio Mariotti, Tony Ukety, Donatella, Pascolini AF, David Friedman, Hugh Taylor, Richard Le Mesurier, Christian Garms, Richard, Porter NR, Abi Smith and Louis Pizzarello. Global Initiative for the Elimination of Avoidable Blindness : action plan 2006-2011. Vision 2020: World Health Organization; 2007.

100. Organization WH. Blindness as a public health problem in China 2016. Available from: http://www.who.int/mediacentre/factsheets/fs230/en/.

101. Seah SKL, Wong TY, Foster PJ, Ng TP, Johnson GJ. Prevalence of lens opacity in Chinese residents of Singapore: the tanjong pagar survey. Ophthalmology.109(11):2058-64.

102. Lin P-Y, Tsai S-Y, Cheng C-Y, Liu J-H, Chou P, Hsu W-M. Prevalence of dry eye among an elderly Chinese population in Taiwan: the Shihpai Eye Study. Ophthalmology. 2003;110(6):1096-101.

103. Wong T, Loon S, Saw S. The epidemiology of age related eye diseases in Asia. British Journal of Ophthalmology. 2006;90(4):506-11.

104. Thylefors B. A GLOBAL INITIATIVE FOR THE ELIMINATION OF AVOIDABLE BLINDNESS. Community Eye Health. 1998;11(25):1-3.

105. Xu J, Zhu S, Li S, Pizzarello L. Models for improving cataract surgical rates in southern China. British journal of ophthalmology. 2002;86(7):723-4.

106. Zhang X, Cotch MF, Ryskulova A, Primo SA, Nair P, Chou C-F, et al. Vision health disparities in the United States by race/ethnicity, education, and economic status: findings from two nationally representative surveys. American journal of ophthalmology. 2012;154(6):S53-S62. e1.

107. Prokofyeva E, Wegener A, Zrenner E. Cataract prevalence and prevention in Europe: a literature review. Acta ophthalmologica. 2013;91(5):395-405.

108. Council TV. PROTECTION FOR THE NAKED EYE: Sunglasses as a Health Necessity. The Vision Council, 2015.

109. Bekker M. Phacoemulsification for cataracts The Encyclopedia of Surgery2016. Available from: http://www.surgeryencyclopedia.com/Pa-St/Phacoemulsification-for-Cataracts.html.

110. Frey R. Extracapsular cataract extraction The Encyclopedia of Surgery2016. Available from: http://www.surgeryencyclopedia.com/Ce-Fi/Extracapsular-Cataract-Extraction.html.

111. Foster A. VISION 2020: THE CATARACT CHALLENGE. Community Eye Health. 2000;13(34):17-9.

112. GoDARTS, Group UDPS, 2 WTCCC. Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes. Nature genetics. 2011;43(2):117-20.

113. Fagerholm E, Ahlqvist E, Forsblom C, Sandholm N, Syreeni A, Parkkonen M, et al. SNP in the genome-wide association study hotspot on chromosome 9p21 confers susceptibility to diabetic nephropathy in type 1 diabetes. Diabetologia. 2012;55(9):2386-93.

114. Facts About Cataract: National Eye Institute; 2009 [updated September]. Available from: https://www.nei.nih.gov/health/cataract/cataract_facts.

115. Cruickshanks KJ, Klein B, Klein R. Ultraviolet light exposure and lens opacities: the Beaver Dam Eye Study. American journal of public health. 1992;82(12):1658-62.

116. WHO. BMI classification 2016. Available from: http://apps.who.int/bmi/index.jsp?introPage=intro_3.html.

117. Klein B, Klein R, Linton K, Franke T. Cigarette smoking and lens opacities: the Beaver Dam Eye Study. American Journal of Preventive Medicine. 1992;9(1):27-30.

118. Flaye D, Sullivan K, Cullinan T, Silver J, Whitelocke R. Cataracts and cigarette smoking: the City Eye Study. Eye. 1989;3(4):379-84.

119. Lea W. UK Chief Medical Officers'Alcohol Guidelines Review. In: Health Do, editor. 2016.

120. Scottish Index of Multiple Deprivation 2016 2016 [updated August 31, 2016]. Available from: http://www.gov.scot/Publications/2016/08/6427.

121. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. The American Journal of Human Genetics. 2007;81(3):559-75.

122. Ellis T. Scotland's Gensus 2011 General Report In: Parliament S, editor. 2015.

123. Nirmalan P, Robin A, Katz J, Tielsch J, Thulasiraj R, Krishnadas R, et al. Risk factors for age related cataract in a rural population of southern India: the Aravind Comprehensive Eye Study. British journal of ophthalmology. 2004;88(8):989-94.

124. dbSNP. Reference SNP (refSNP) Cluster Report: rs10197646 NCBI2016. Available from: http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=10197646.

125. Consortium GP. A map of human genome variation from population-scale sequencing. Nature. 2010;467(7319):1061-73.

126. NCBI. Reference SNP (refSNP) Cluster Report: rs7582173 NCBI2016. Available from: http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=7582173.

127. NCBI. Reference SNP (refSNP) Cluster Report: rs62168795 NCBI2016. Available from: http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=rs62168795.

128. NCBI. Reference SNP (refSNP) Cluster Report: rs1381015 2016. Available from: http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=rs1381015.

129. NCBI. Reference SNP (refSNP) Cluster Report: rs2269547 NCBI; 2016. Available from: http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=rs2269547.

130. NCBI. Reference SNP (refSNP) Cluster Report: rs523355 NCBI; 2016. Available from: http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=rs523355.

131. NCBI. CCT7 chaperonin containing TCP1 subunit 7 [Homo sapiens (human)] 2016. Available from: http://www.ncbi.nlm.nih.gov/gene?cmd=Retrieve&dopt=Graphics&list_uids=10574.

132. Pearson G, Robinson F, Beers Gibson T, Xu B-e, Karandikar M, Berman K, et al. Mitogen-activated protein (MAP) kinase pathways: regulation and physiological functions 1. Endocrine reviews. 2001;22(2):153-83.

133. Li MJ, Wang P, Liu X, Lim EL, Wang Z, Yeager M, et al. GWASdb: a database for human genetic variants identified by genome-wide association studies. Nucleic acids research. 2011:gkr1182.

134. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proceedings of the National Academy of Sciences. 2009;106(23):9362-7.

135. Lab J. GWAS information on cataract 2016. Available from: http://jjwanglab.org/gwasdb/gwasdb2/gwasdb2/go_trait/DOID:83.

136. Lab J. GWAS information on diabetes mellitus 2016. Available from: http://jjwanglab.org/gwasdb/gwasdb2/gwasdb2/go_trait/DOID:9351.

137. KINOSHITA JH. Mechanisms initiating cataract formation proctor lecture. Investigative Ophthalmology & Visual Science. 1974;13(10):713-24.

138. Kinoshita JH, Fukushi S, Kador P, Merola LO. Aldose reductase in diabetic complications of the eye. Metabolism. 1979;28(4):462-9.

139. Takamura Y, Sugimoto Y, Kubo E, Takahashi Y, Akagi Y. Immunohistochemical study of apoptosis of lens epithelial cells in human and diabetic rat cataracts. Japanese journal of ophthalmology. 2001;45(6):559-63.

140. Li W-C, Kuszak JR, Dunn K, Wang R-R, Ma W, Wang G-M, et al. Lens epithelial cell apoptosis appears to be a common cellular basis for non-congenital cataract development in humans and animals. The Journal of cell biology. 1995;130(1):169-81.

141. Lab J. GWAS information on hypertension 2016. Available from: http://jjwanglab.org/gwasdb/gwasdb2/gwasdb2/go_trait/DOID:10763.

142. Leske MC, Wu S-Y, Hennis A, Connell AM, Hyman L, Schachat A, et al. Diabetes, hypertension, and central obesity as cataract risk factors in a black population: The Barbados Eye Study11Members of the Barbados Eye Study Group are listed in the Appendix at the end of this article. Ophthalmology. 1999;106(1):35-41.

143. Naish JC, Denise Syndercombe. Medical Sciences. 2nd ed2014.

144. Poulter NR, Prabhakaran D, Caulfield M. Hypertension. The Lancet.386(9995):801-12.

145. Carretero OA, Oparil S. Essential Hypertension. Part I: Definition and Etiology. 2000;101(3):329-35.

146. Kibbe WA, Arze C, Felix V, Mitraka E, Bolton E, Fu G, et al. Disease Ontology 2015 update: an expanded and updated database of human diseases for linking biomedical knowledge through disease data. Nucleic acids research. 2014:gku1011.

147. Institute NE. Facts About Myopia 2016. Available from: https://nei.nih.gov/health/errors/myopia.

148. Saw S-M, Katz J, Schein OD, Chew S-J, Chan T-K. Epidemiology of myopia. Epidemiologic reviews. 1996;18(2):175-87.

149. Praveen MR, Vasavada AR, Jani UD, Trivedi RH, Choudhary PK. Prevalence of cataract type in relation to axial length in subjects with high myopia and emmetropia in an Indian population. American journal of ophthalmology. 2008;145(1):176-81. e1.

150. Younan C, Mitchell P, Cumming RG, Rochtchina E, Wang JJ. Myopia and incident cataract and cataract surgery: the Blue Mountains Eye Study. Investigative ophthalmology & visual science. 2002;43(12):3625-32.

151. Wu S-Y, Nemesure B, Leske MC. Refractive errors in a black adult population: the Barbados Eye Study. Investigative ophthalmology & visual science. 1999;40(10):2179-84.

152. McCarty CA, Mukesh B, Fu CL, Taylor HR. The epidemiology of cataract in Australia. American journal of ophthalmology. 1999;128(4):446-65.

153. Institute NE. Facts About Age-Related Macular Degeneration: NIH; 2016. Available from: https://nei.nih.gov/health/maculardegen/armd_facts.

154. Mehta S. Age-Related Macular Degeneration. Primary Care: Clinics in Office Practice.42(3):377-91.

155. Vos T, Barber RM, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet.386(9995):743-800.

156. Freeman EE, Munoz B, West SK, Tielsch JM, Schein OD. Is there an association between cataract surgery and age-related macular degeneration? Data from three population-based studies. American journal of ophthalmology. 2003;135(6):849-56.

157. NCBI. TCP1 t-complex 1. 2016.

158. Knee KM, Sergeeva OA, King JA. Human TRiC complex purified from HeLa cells contains all eight CCT subunits and is active in vitro. Cell Stress and Chaperones. 2013;18(2):137-44.

159. Sergeeva OA. Assembly and substrate recognition properties of human CCT subunits of the TRiC chaperonin: Massachusetts Institute of Technology; 2014.

160. Lamb TD, Collin SP, Pugh EN. Evolution of the vertebrate eye: opsins, photoreceptors, retina and eye cup. Nature Reviews Neuroscience. 2007;8(12):960-76.

161. Andley UP. The lens epithelium: Focus on the expression and function of the α-crystallin chaperones. The international journal of biochemistry & cell biology. 2008;40(3):317-23.

162. Andley UP, Song Z, Wawrousek EF, Bassnett S. The molecular chaperone αA-crystallin enhances lens

epithelial cell growth and resistance to UVA stress. Journal of Biological Chemistry. 1998;273(47):31252-61.

163. Cherian M, Abraham E. Decreased molecular chaperone property of α-crystallins due to posttranslational modifications. Biochemical and biophysical research communications. 1995;208(2):675-9.

164. NCBI. R3HDM1 R3H domain containing 1 2016. Available from: http://www.ncbi.nlm.nih.gov/gene?cmd=Retrieve&dopt=Graphics&list_uids=23518.

165. Project E. Gene: R3HDM1 2016. Available from: http://www.ensembl.org/Homo_sapiens/Gene/ExpressionAtlas?g=ENSG00000048991;r=2:135531455-135725270.

166. GeneCards. R3HDM1 Gene: GeneCards; 2016. Available from: http://www.genecards.org/cgi-bin/carddisp.pl?id_type=hgnc&id=9757.

167. Goldmann H. Senile changes of the lens and the vitreous: the Arthur J. Bedell lecture. American journal of ophthalmology. 1964;57(1):1-13.

168. LAURSEN AB, Fledelius H. Variations of lens thickness in relation to biomicroscopic types of human senile cataract. Acta ophthalmologica. 1979;57(1):1-13.

169. Karim A, Jacob T, Thompson G. The human anterior lens capsule: cell density, morphology and mitotic index in normal and cataractous lenses. Experimental eye research. 1987;45(6):865-74.

170. Konofsky K, Naumann GO, Guggenmoos-Holzmann I. Cell density and sex chromatin in lens epithelium of human cataracts: quantitative studies in flat preparation. Ophthalmology. 1987;94(7):875-80.

171. Harding JJ, Crabbe MJC. The lens: development, proteins, metabolism and cataract. The Eye, Davson, H(Ed). 1984;1.

172. David LL, Azuma M, Shearer TR. Cataract and the acceleration of calpain-induced beta-crystallin insolubilization occurring during normal maturation of rat lens. Investigative ophthalmology & visual science. 1994;35(3):785-93.

173. Shearer TR, David LL, Anderson RS, Azuma M. Review of selenite cataract. Current eye research. 1992;11(4):357-69.

174. Iwasaki N, David LL, Shearer TR. Crystallin degradation and insolubilization in regions of young rat lens with calcium ionophore cataract. Investigative ophthalmology & visual science. 1995;36(2):502-9.

175. NCBI. GGA1 golgi associated, gamma adaptin ear containing, ARF binding protein 1: NCBI; 2016. Available from: http://www.ncbi.nlm.nih.gov/gene?cmd=Retrieve&dopt=Graphics&list_uids=26088.

176. Doray B, Ghosh P, Griffith J, Geuze HJ, Kornfeld S. Cooperation of GGAs and AP-1 in packaging MPRs at the trans-Golgi network. Science. 2002;297(5587):1700-3.

177. Ginger RS, Askew SE, Ogborne RM, Wilson S, Ferdinando D, Dadd T, et al. SLC24A5 encodes a trans-Golgi network protein with potassium-dependent sodium-calcium exchange activity that regulates human epidermal melanogenesis. Journal of Biological Chemistry. 2008;283(9):5486-95.

178. Dolman NJ, Tepikin AV. Calcium gradients and the Golgi. Cell calcium. 2006;40(5):505-12.

179. Maddala R, Nagendran T, de Ridder GG, Schey KL, Rao PV. L-type calcium channels play a critical role in maintaining lens transparency by regulating phosphorylation of aquaporin-0 and myosin light chain and expression of connexins. PloS one. 2013;8(5):e64676.

180. Meissner A, Noack T. Proliferation of human lens epithelial cells (HLE-B3) is inhibited by blocking of voltage-gated calcium channels. Pflügers Archiv-European Journal of Physiology. 2008;457(1):47-59.

181. Tang D, Borchman D, Yappert MC, Vrensen GF, Rasi V. Influence of age, diabetes, and cataract on calcium, lipid-calcium, and protein-calcium relationships in human lenses. Investigative ophthalmology &

visual science. 2003;44(5):2059-66.
182. Vision Problems in the U.S. Estimated Age-Specific Prevalence Rates for Cataract Vision Problems in the U.S.2012. Available from: http://www.visionproblemsus.org/cataract.html.
183. Klein BE, Klein R, Lee KE. Incidence of age-related cataract: the Beaver Dam Eye Study. Archives of Ophthalmology. 1998;116(2):219-25.
184. Vaughan D, Cook RD, Asbury T. General ophthalmology: Lange Medical Publications; 1971.
185. Ivers RQ, Cumming RG, Mitchell P, Peduto AJ. Diabetes and risk of fracture the blue mountains eye study. Diabetes care. 2001;24(7):1198-203.
186. Waudby CJ, Berg RL, Linneman JG, Rasmussen LV, Peissig PL, Chen L, et al. Cataract research using electronic health records. BMC ophthalmology. 2011;11(1):1.
187. Schwab IR, Dawson CR, Hoshiwara I, Szuter CF, Knowler WC. Incidence of cataract extraction in Pima Indians: diabetes as a risk factor. Archives of Ophthalmology. 1985;103(2):208-12.
188. Janghorbani M, Jones RB, Allison SP. Incidence of and risk factors for cataract among diabetes clinic attenders. Ophthalmic Epidemiology. 2000;7(1):13-25.
189. Health UDo, Services H. The biology and behavioral basis for smoking-attributable disease: a report of the surgeon general. Atlanta: US Department of Health and Human Services, Centers for Disease Control and Prevention. National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. 2010.
190. Zhang X, Kahende J, Fan AZ, Barker L, Thompson TJ, Mokdad AH, et al. Smoking and visual impairment among older adults with age-related eye diseases. Prev Chronic Dis. 2011;8(4):A84.
191. Saliba A. Impact of rurality on optical health: review of the literature and relevant Australian Bureau of Statistics data. Rural and remote health. 2008;8(1056).
192. Lindblad BE, Håkansson N, Wolk A. Smoking cessation and the risk of cataract: a prospective cohort study of cataract extraction among men. JAMA ophthalmology. 2014;132(3):253-7.
193. Vision RY. Alcohol, Cataracts, and Other Drinking-Related Visual Disorders: Rebuild Your Vision; 2016. Available from: http://www.rebuildyourvision.com/blog/food-for-your-eyes/alcohol-cataracts-and-other-drinking-related-visual-disorders/.
194. Cumming RG, Mitchell P. Alcohol, smoking, and cataracts: the Blue Mountains eye study. Archives of Ophthalmology. 1997;115(10):1296-303.
195. Kuang T-M, Tsai S-Y, Hsu W-M, Cheng C-Y, Liu J-H, Chou P. Body mass index and age-related cataract: the Shihpai Eye Study. Archives of Ophthalmology. 2005;123(8):1109-14.
196. Hiller R, Podgor MJ, Sperduto RD, Nowroozi L, Wilson PW, D'Agostino RB, et al. A longitudinal study of body mass index and lens opacities: the Framingham Studies. Ophthalmology. 1998;105(7):1244-50.
197. Vrijheid M, Dolk H, Stone D, Abramsky L, Alberman E, Scott J. Socioeconomic inequalities in risk of congenital anomaly. Archives of Disease in Childhood. 2000;82(5):349-52.
198. Ho VH, Schwab IR. Social economic development in the prevention of global blindness. British Journal of Ophthalmology. 2001;85(6):653-7.
199. Marmot M. Social determinants of health inequalities. The Lancet. 2005;365(9464):1099-104.
200. Office M. Eastern Scotland: climate 2015. Available from: http://www.metoffice.gov.uk/climate/uk/regional-climates/es.
201. Schaumberg DA, Glynn RJ, Christen WG, Ajani UA, Stürmer T, Hennekens CH. A prospective study of blood pressure and risk of cataract in men. Annals of epidemiology. 2001;11(2):104-10.
202. Goldacre MJ, Wotton CJ, Keenan TD. Risk of selected eye diseases in people admitted to hospital for hypertension or diabetes mellitus: record linkage studies. British journal of ophthalmology. 2012:bjophthalmol-2012-301519.

203. Smith LL. Review of progress in sterol oxidations: 1987–1995. Lipids. 1996;31(5):453-87.
204. Girao H, Mota C, Pereira P. Cholesterol may act as an antioxidant in lens membranes. Current eye research. 1999;18(6):448-54.
205. Meyer D, Parkin D, Maritz F, Liebenberg P. Abnormal serum lipoprotein levels as a risk factor for the development of human lenticular opacities. 2003.
206. Johnson CY. Boston scientists say triglycerides play key role in heart health The Boston Globe2014. Available from: http://www.bostonglobe.com/news/science/2014/06/18/boston-researchers-find-that-triglycerides-play-pivotal-role-heart-health/ynrM4QQwIq1fCCoRwMfOAN/story.html.
207. Drummond KE, Brefere LM. Nutrition for foodservice and culinary professionals: J. Wiley; 2004.
208. Husain R, Tong L, Fong A, Cheng JF, How A, Chua W-H, et al. Prevalence of cataract in rural Indonesia. Ophthalmology. 2005;112(7):1255-62.
209. Stocks N, Patel R, Sparrow J, Davey-Smith G. Prevalence of cataract in the Speedwell Cardiovascular Study: a cross-sectional survey of men aged 65–83. Eye. 2002;16(3):275-80.
210. Gibson J, Rosenthal A, Lavery J. A study of the prevalence of eye disease in the elderly in an English community. Transactions of the ophthalmological societies of the United Kingdom. 1984;104:196-203.
211. Frost A, Hopper C, Frankel S, Peters TJ, Durant J, Sparrow J. The population requirement for cataract extraction: a cross-sectional study. Eye. 2001;15(6):745-52.
212. Evans J, Fletcher A, Wormald R. Causes of visual impairment in people aged 75 years and older in Britain: an add-on study to the MRC Trial of Assessment and Management of Older People in the Community. British Journal of Ophthalmology. 2004;88(3):365-70.

Appendix

List of Abbreviated Word

SNP: Single Nucleotide Polymorphisms

BMI: Body max index

WHO: World Health Organization

GWAS: Genome-Wide Association Studies

LOCS: Lens Opacity Classification System

AR: aldose reductase

STZ: Streptozotocin

AGE: Advanced glycation end

GSH: Glutathione

TCF7L2: Transcription Factor 7-like 2

RR: Relative risk

UV-B: Ultraviolet B

IOL: Intraocular lens implant

ICCE: Intracapsular cataract extraction

ECCE: extra capsular cataract extraction

NHS: National Health Service

GoDARTs: Genetics of Diabetes Audit and Research Tayside

SCI-DC: Scottish Care Information-Diabetes Collaboration

WTCCC2 Wellcome Trust Case Control Consortium 2

SUMMIT: Surrogate markers for Micro- and Macro-vascular hard endpoints for Innovative diabetes Tools

NHANES: National Health and Nutrition Examination Survey

SIMD: Scottish index of multiple deprivation

AMD: Age-related macular degeneration